

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:12:44 ; Search time 200 Seconds
(without alignments)
2926.263 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGALSNLFPNNGROF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

Database :

A_Geneseq_21.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*
9: Geneseq2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	8	ADJ27297
2	6909	100.0	1332	9	ADY60740 Human NPC
3	6909	100.0	1332	9	AEB93570 Human NPC
4	6896	99.8	1332	4	AAM79169 Human pro
5	6896	99.8	1332	4	AAG65638 SSD-conta
6	6872.5	99.5	1359	4	AAM79168 Human pro
7	6872.5	99.5	1359	8	ADJ27337 Human NPC
8	6872.5	99.5	1359	9	ADY60780 Human NPC
9	6872.5	99.5	1359	9	AEB93610 Human NPC
10	6536	94.6	1344	4	ABG22693 Novel hum
11	5421.5	78.5	1331	8	ADJ27295 Rat NPCIL
12	5421.5	78.5	1331	9	ADY60738 Rat Niema
13	5421.5	78.5	1331	9	AEB93568 Rat NPCIL
14	5407	78.3	1333	8	ADJ27305 Mouse NPC
15	5407	78.3	1333	9	ADY60748 Mouse Nie
16	5407	78.3	1333	9	AEB93578 Mouse NPC
17	4466	64.6	982	4	ABG22691 Novel hum
18	2402.5	34.8	1278	3	AAB88445 Human NPC
19	2402.5	34.8	1278	3	AAB842983 Human ORF
20	2402.5	34.8	1278	8	ADQ39879 Human myo
21	2402.5	34.8	1278	8	ADU06723 Novel bro
22	2385	34.5	1319	2	AAM88446 Mouse NPC
23	2262	32.7	1287	4	AB61737 Drosophil
24	2262	32.7	1287	8	ADS96670 Drosophil

ALIGNMENTS

RESULT 1

ADJ27297

ID ADJ27297 standard; protein; 1332 AA.

XX AC ADJ27297;

XX DT 20-MAY-2004 (first entry)

XX DE Human NPCIL1.

XX KW Niemann-Pick disease; type C1; gene-like 1; NPCIL1; trans-golgi network;

XX KW plasma membrane; transport signal; promoter;

XX KW sterol regulated element binding protein 1; SREBP1;

XX KW binding consensus sequence; transmembrane domain; sterol-sensing domain;

XX KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;

XX KW cholesterol absorption; serum cholesterol; hyperlipidaemia;

XX KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX OS Homo sapiens.

XX PN WO2004009772-A2.

XX PD 29-JAN-2004.

XX PF 17-JUL-2003; 2003WO-US022467.

XX PR 19-JUL-2002; 2002US-0397442P.

XX PI (SCHE) SCHERING CORP.

XX PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX DR WPI; 2004-132945/13.

XX DR N-PSDB; ADJ27296.

XX XX GENBANK; AF192522.

PT New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting or identifying antagonists of NPCIL1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

PS Claim 11; SEQ ID NO 4; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1 (NPCIL1) polypeptide. NPCIL1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and

Abb58629 Drosophil
Aaw88447 Yeast NPC
Aaw89448 Caenorhab
Aam80153 Human pro
Aam80152 Human pro
Aag65637 SSD domal
Aar75375 Human pat
Aaw52200 Human pat
Aaw72969 Human pat
Aab67163 Human pat
Aae19830 Human pat
Abj10931 TRC8 rela
Abu62275 Human pat
Add46678 Human pro
Ade94224 Human ptc
Adh62731 Human pat
Ade48989 Human pat
Aaw52199 Mouse pat
Aaw72968 Mouse pat
Aab67159 Murine pa

CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels,
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42 amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.

XX
SQ Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 8; Length 1332;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSPPYTHIOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDVDSRVRVPA 180
DB 121 VNLHCHNTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFABQSDVDSRVRVPA 180
QY 181 ATLAAGTGVGVSALCNQRMWLFQDGTGNGLAPLDTIFHLEPGQAVGSGIQLPNEG 240
DB 181 ATLAAGTGVGVSALCNQRMWLFQDGTGNGLAPLDTIFHLEPGQAVGSGIQLPNEG 240
QY 241 ARCNEQSGDDVATCSQDCQCAASPAJARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNEQSGDDVATCSQDCQCAASPAJARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKWDPKGTSLSDKLSFSFHTLLGQFFQCGWGTWVASWELTILVLSV 360
DB 301 LLVGFVAPARDKSKWDPKGTSLSDKLSFSFHTLLGQFFQCGWGTWVASWELTILVLSV 360
QY 361 IPVVALAAGLVFTLTDTPVELMSAPNSQARSKA FHDQHFGRFTNQVILTPNRSY 420
DB 361 IPVVALAAGLVFTLTDTPVELMSAPNSQARSKA FHDQHFGRFTNQVILTPNRSY 420
QY 421 RYDSSLGPKNFGSILDLIDLLELLEQLERLRLQVMSPEAQARNISLQDICYAPLNPDT 480
DB 421 RYDSSLGPKNFGSILDLIDLLELLEQLERLRLQVMSPEAQARNISLQDICYAPLNPDT 480
QY 481 SLVDCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540
DB 481 SLVDCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCWADYCAPVPFLATGGYKGDYSRAELIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
DB 541 SCWADYCAPVPFLATGGYKGDYSRAELIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
QY 601 RAFQRMAGNFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSWSRV 660
DB 601 RAFQRMAGNFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSWSRV 660
QY 661 MVDSKATLGLGVAVVLGAVMAAGPFSYLGIRSSILVQVPLVLSVADNIFIFVLE 720
DB 661 MVDSKATLGLGVAVVLGAVMAAGPFSYLGIRSSILVQVPLVLSVADNIFIFVLE 720
QY 721 YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFGLALTTPPAVRTFALTSLAV 780
DB 721 YQRLPRPGPREVHIGRALGRVAPSMMLCSLSEALCFGLALTTPPAVRTFALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840

DB 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPQELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVWLLFLALFGVSLYSYCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVWLLFLALFGVSLYSYCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTGLYGFSEAGMNAICSSAGCANNFSTQKIYATEFPPEQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTGLYGFSEAGMNAICSSAGCANNFSTQKIYATEFPPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSYEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSYEQPHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQOVLASRFMAYHKLNQSDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQOVLASRFMAYHKLNQSDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEVFPYTIITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEVFPYTIITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSTVMLVDTVGFMAWDISYNAVSLINLVSAVGMSEVFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSTVMLVDTVGFMAWDISYNAVSLINLVSAVGMSEVFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFRNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNITYVNHSPFGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNITYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332
RESULT 2
ADY60740
ID ADY60740 standard; protein; 1332 AA.
XX
AC ADY60740;
XT 19-MAY-2005 (first entry)
XX Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 4.
XX Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX Homo sapiens.
XX WO2005015988-A1.
XX 24-FEB-2005.
XX 16-DEC-2003; 2003WO-US040113.
XX 17-JUL-2003; 2003US-00621758.
XX 22-AUG-2003; 2003US-00646301.
XX 16-SEP-2003; 2003US-00663208.
XX (SCHE) SCHERING CORP.
XX Altman SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
XX N-PSDB; ADY60739, ADY60787.
DB

AC	AAM79169;	
XX	06-NOV-2001 (first entry)	
DT	Human protein SEQ ID NO 1831.	
XX		
DE		
XX	Human; cytokine; cell proliferation; cell differentiation; gene therapy;	
KW	vaccine; peptide therapy; stem cell growth factor; haematopoiesis;	
KW	tissue growth factor; immunomodulatory; cancer; leukaemia;	
KW	nervous system disorder; arthritis; inflammation.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200157190-A2.	
XX		
PD	09-AUG-2001.	
XX		
PF	05-FEB-2001; 2001WO-US004098.	
XX		
PR	03-FEB-2000; 2000US-00496914.	
PR	27-APR-2000; 2000US-00560875.	
PR	20-JUN-2000; 2000US-00598075.	
PR	19-JUL-2000; 2000US-00620325.	
PR	01-SEP-2000; 2000US-00654936.	
PR	15-SEP-2000; 2000US-00663561.	
PR	20-OCT-2000; 2000US-00693325.	
PR	30-NOV-2000; 2000US-00728422.	
XX		
PA	(HYSE-) HYSEQ INC.	
XX		
PI	Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	
PI	Ma Y, Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	
XX		
DR	WPI; 2001-476283/51.	
DR	N-PSDB; AAK52302.	
XX		
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	
PT	in diagnosis and gene therapy.	
PS	Claim 20; Page 4216-4219; 6221pp; English.	
XX		
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the	
CC	encoded polypeptides (AAM79323-AAW80302) that exhibit activity elating to	
CC	cytokine, cell proliferation or cell differentiation or which may induce	
CC	production of other cytokines in other cell populations. The	
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	
CC	peptide therapy. The polypeptides have various cytokine-like activities,	
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	
CC	activity, tissue growth factor activity, immunomodulatory activity and	
CC	activin/inhibin activity and may be useful in the diagnosis and/or	
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	
CC	inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111	
CC	(AAK52582) and 3666 (AAW80020) are omitted as the relevant pages from the	
CC	sequence listing were missing at the time of publication	
XX		
SQ	Sequence 1332 AA;	
	Query Match 99.8%; Score 6896; DB 4; Length 1332;	
	Best Local Similarity 99.8%; Pred. No. 0;	
	Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	1 MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60	
DB	1 MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60	
QY	61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120	
DB	61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120	
QY	121 VNLHCHNTCSNQSIFINVTVAQLGAGQLPAVVAEAFYQHSFABQSYDCSRVRPAA 180	
DB	121 VNLHCHNTCSNQSIFINVTVAQLGAGQLPAVVAEAFYQHSFABQSYDCSRVRPAA 180	

QY	181 ATLA VGTMC GVGYSALCNAQRWLNFGQDGTGNGLAPLDITFHLLPEQQA VSGGIQPLNEGV 240	
DB	181 ATLA VGTMC GVGYSALCNAQRWLNFGQDGTGNGLAPLDITFHLLPEQQA VSGGIQPLNEGV 240	
QY	241 ARCNE SGGDDVATCSQDCCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	
DB	241 ARCNE SGGDDVATCSQDCCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300	
QY	301 LLVGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQFPQGGWGTWASWPLTILVLSV 360	
DB	301 LLVGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQFPQGGWGTWASWPLTILVLSV 360	
QY	361 IPPVAAAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPPFRFTNVTIAPNRSY 420	
DB	361 IPPVAAAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPPFRFTNVTIAPNRSY 420	
QY	421 RYDSLILGPKPFGSGLDLDLLELLELQERLHLQWSPQARNSLQDICYAPLNPDT 480	
DB	421 RYDSLILGPKPFGSGLDLDLLELLELQERLHLQWSPQARNSLQDICYAPLNPDT 480	
QY	481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDMKDHFLYCANAPLT PKDGTALAL 540	
DB	481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDMKDHFLYCANAPLT PKDGTALAL 540	
QY	541 SCWADYGAPVPPFLAIGGYKGYSEAEALIMTFSILNNYPAGDPRLAQA KLEEAFLBEM 600	
DB	541 SCWADYGAPVPPFLAIGGYKGYSEAEALIMTFSILNNYPAGDPRLAQA KLEEAFLBEM 600	
QY	601 RAPQRMAGMQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISIALSGYSWSRV 660	
DB	601 RAPQRMAGMQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISIALSGYSWSRV 660	
QY	661 MYDSKATILGCGVAVVLGAVMAAMGFPFSLGTRSSLVILQVVPFLVLSV GADNIFIVLE 720	
DB	661 MYDSKATILGCGVAVVLGAVMAAMGFPFSLGTRSSLVILQVVPFLVLSV GADNIFIVLE 720	
QY	721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAICFFLGLALT PMPAVRTALTSGLA V 780	
DB	721 YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEAICFFLGLALT PMPAVRTALTSGLA V 780	
QY	781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPGQGGGLLPGFPQKAYAPPL 840	
DB	781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKPQELPPPGQGGGLLPGFPQKAYAPPL 840	
QY	841 LHMITRGVLLPLALFGVSLYSMCHISVGLDQELALPKDSVILLDYFLPLNRYFVGAPV 900	
DB	841 LHMITRGVLLPLALFGVSLYSMCHISVGLDQELALPKDSVILLDYFLPLNRYFVGAPV 900	
QY	901 YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFDIW 960	
DB	901 YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFDIW 960	
QY	961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKLVPFLNDRP 1020	
DB	961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHMSITMGSVRPSVEQFHKLVPFLNDRP 1020	
QY	1021 NIKCPKGLAAYSTSVNLTSDQVLA SRPMAYHKLKNSQDYTEALRAARELANITADL 1080	
DB	1021 NIKCPKGLAAYSTSVNLTSDQVLA SRPMAYHKLKNSQDYTEALRAARELANITADL 1080	
QY	1081 RKVPGTDPAPVPPVPTTINVFYEOVLTLPGLFMLSCLVPTTAVSCLLGLDLSGLL 1140	
DB	1081 RKVPGTDPAPVPPVPTTINVFYEOVLTLPGLFMLSCLVPTTAVSCLLGLDLSGLL 1140	
QY	1141 NLLSIVMLVDTVGFMALWDISYNVAVSLINIVSAGMSVEFVSHITRSPFAISTKPTWLER 1200	
DB	1141 NLLSIVMLVDTVGFMALWDISYNVAVSLINIVSAGMSVEFVSHITRSPFAISTKPTWLER 1200	
QY	1201 AKEATISNGSAVPAGVAMTNLPGILVLGLAKAQILQIOPFFRLNLLITLLGLHLGVLPV 1260	
DB	1201 AKEATISNGSAVPAGVAMTNLPGILVLGLAKAQILQIOPFFRLNLLITLLGLHLGVLPV 1260	

Qy	1261	ILSYVGPVNPALALEQKRAEEAAVAVVASCPNHPSRVSTADNIYNHSHFEGSIKAGA	1320
Db	1261	ILSYVGPVNPALALEQKRAEEAAVAVVASCPNHPSRVSTADNIYNHSHFEGSIKAGA	1320
Qy	1321	ISNPLPNNGRQF	1332
Db	1321	ISNPLPNNGRQF	1332
RESULT 5			
ID	AAG65638	standard; protein; 1332 AA.	
XX			
AC	AAG65638;		
XX			
DT	07-JAN-2002	(first entry)	
DE			
XX			
XX		SSD-containing SSP1 protein sequence.	
KW	SSD; sterol-sensing domain; human; liver; testis; brain; cancer;		
KW	neotropic; neuroprotective; antidiabetic; antiarteriosclerotic;		
KW	cytostatic; antilipemic; SSP1.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200170974-A1.		
XX			
PD	27-SEP-2001.		
XX			
PF	22-MAR-2001; 2001WO-JP002279.		
XX			
PR	24-MAR-2000; 2000JP-00088595.		
XX			
PA	(TAKE) TAKEDA CHEM IND LTD.		
XX			
PI	Taniyama Y, Kita S, Komiya T;		
XX			
DR	WPI; 2001-611501/70.		
DR	N-PSDB; AAI66914.		
XX			
PT	New sterol-sensing domain-containing protein for diagnosing and screening		
PT	candidate compounds in drug development for diabetes, obesity, cancer,		
PT	arteriosclerosis, hyperlipidemia and neurodegenerative disorders.		
XX			
PS	Claim 7; Page 115-122; 171pp; Japanese.		
XX			
CC	The invention provides a novel SSD (sterol-sensing domain)-containing		
CC	protein. The protein originates from human liver, human testis or human		
CC	brain. The protein can be expressed by standard recombinant methodology.		
CC	The proteins, encoded DNAs and antibodies are useful in diagnosis and		
CC	screening candidate compounds in drug development for diabetes, obesity,		
CC	cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such		
CC	as Alzheimer's disease and neural disorders. The present sequence		
CC	represents the SSP1 protein which contains the SSD domain		
XX			
SQ	Sequence 1332 AA;		
Query Match 99.8%; Score 6896; DB 4; Length 1332;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAEAGRGWLLWALLLLRLAQSEPYTTIHQPGYCAFYDECCKNPELSGLMTLSNVCLSN	60
Db	1	MAEAGRGWLLWALLLLRLAQSEPYTTIHQPGYCAFYDECCKNPELSGLMTLSNVCLSN	60
Qy	61	TPARKITGDHILLOKICPRLYTPNTQACCSAKQLVLEASLITKALLTRCPACSDNF	120
Db	61	TPARKITGDHILLOKICPRLYTPNTQACCSAKQLVLEASLITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSNQSLFNVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDCSRRVPAA	180
Db	121	VNLHCHNTCSNQSLFNVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDCSRRVPAA	180

Qy	181	ATLAVGTCGVYGSALCNAORWLNFGQDTGNGIAPLDITFHLLERQQAQVSGIQPLNEGV	240
Db	181	ATLAVGTCGVYGSALCNAORWLNFGQDTGNGIAPLDITFHLLERQQAQVSGIQPLNEGV	240
Qy	241	ARCNESQGDVATCSCQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Db	241	ARCNESQGDVATCSCQDCAASCPAARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Qy	301	LLVGFVAPAROKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV	360
Db	301	LLVGFVAPAROKSKWDPKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHFQFFFTNQVILTPAPNRSSY	420
Db	361	IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDQHFQFFFTNQVILTPAPNRSSY	420
Qy	421	RYDLSLLGPKNFGSILDLLELLELQBLRHLQVWSPQAQRNISILODICYAPLNPDT	480
Db	421	RYDLSLLGPKNFGSILDLLELLELQBLRHLQVWSPQAQRNISILODICYAPLNPDT	480
Qy	481	SLYDCCINSILQYFQNNRTLLLTANQTLWGQTSQVDWKDHFLYCANAPLTFKDGATLAL	540
Db	481	SLYDCCINSILQYFQNNRTLLLTANQTLWGQTSQVDWKDHFLYCANAPLTFKDGATLAL	540
Qy	541	SCNADYGAPVFPFLAIGGYKGYSEAEALIMTFSLMNYPAGDPRLAQAKLWBEAFLEEM	600
Db	541	SCNADYGAPVFPFLAIGGYKGYSEAEALIMTFSLMNYPAGDPRLAQAKLWBEAFLEEM	600
Qy	601	RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSVSSSRV	660
Db	601	RAFQRMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSVSSSRV	660
Qy	661	MVDSKATILGGVAVVLTGAVMAAGFESYLGIRSSVLVQVPPVLVSVGADNIFIFVLE	720
Db	661	MVDSKATILGGVAVVLTGAVMAAGFESYLGIRSSVLVQVPPVLVSVGADNIFIFVLE	720
Qy	721	YQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGCALTPMPAPVTFALTSLAV	780
Db	721	YQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGCALTPMPAPVTFALTSLAV	780
Qy	781	ILDFLQMSAFVALLSDSKRQEARLDVCCVQKQELPPPGQEGGLLGPQKAYAPPL	840
Db	781	ILDFLQMSAFVALLSDSKRQEARLDVCCVQKQELPPPGQEGGLLGPQKAYAPPL	840
Qy	841	LHWITRGVVLTLFALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGA	900
Db	841	LHWITRGVVLTLFALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGA	900
Qy	901	YFVTLTGYNFSSEAGMNAICSSAGCNPSPFTQKIQYATEPPEQSYLAIPASSWDDFDW	960
Db	901	YFVTLTGYNFSSEAGMNAICSSAGCNPSPFTQKIQYATEPPEQSYLAIPASSWDDFDW	960
Qy	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNKSITWGSVRPSVEQPHKYLFWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNKSITWGSVRPSVEQPHKYLFWFLNDRP	1020
Qy	1021	NLKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPKNSQDYTEALRAARELAANITADL	1080
Db	1021	NLKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPKNSQDYTEALRAARELAANITADL	1080
Qy	1081	RKVPGTDPAPFEPYPTITNVFYEQYLTILPEGLFMLSICLVPTPAVSCLLGLDLSGLL	1140
Db	1081	RKVPGTDPAPFEPYPTITNVFYEQYLTILPEGLFMLSICLVPTPAVSCLLGLDLSGLL	1140
Qy	1141	NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER	1200
Db	1141	NLLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER	1200
Qy	1201	AKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLHLGLVFLPV	1260
Db	1201	AKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLHLGLVFLPV	1260
Qy	1261	ILSYVGPVNPALALEQKRAEEAAVAVVASCPNHPSRVSTADNIYNHSHFEGSIKAGA	1320

Db 1261 ILSYGVDPNPALEQKRAEAVAAVMVASCNPHSRVSTADNYYVNSHFGSKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
RESULT 6
AAM79168
ID AAM79168 standard; protein; 1359 AA.
XX
AC AAM79168;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1830.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
FN WO200157190-A2.
XX
PD
XX
XX 09-AUG-2001.
XX
XX 05-FEB-2001; 2001WO-US004098.
XX
XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52301.
XX
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
XX Claim 20; Page 4213-4216; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
XX sequence listing were missing at the time of publication
SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKI CPRLTYGNTQACCSAKQLVSLASLSTYKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKI CPRLTYGNTQACCSAKQLVSLASLSTYKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSILFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSILFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMGVYGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCSCQDCMAASCPAIAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSCQDCMAASCPAIAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVRVAPARDKSKWVDPKKGTSLSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKWVDPKKGTSLSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPFRFTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSQARSEKAFHQHFGPFRFTNQVILTAPNRSY 420
QY 421 RYDSLLLGPKNPSGILDLDLLELLELQERLHLQVWSPQAQRNLSQDICYAPLNPDNT 480
Db 421 RYDSLLLGPKNPSGILDLDLLELLELQERLHLQVWSPQAQRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFONNETLLLTANOTLMGQTSQVDWKDHELYCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFONNETLLLTANOTLMGQTSQVDWKDHELYCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAPQRMAGMFQVTPAERSLEDEINRTTABDLPIFATSYIVIFLYISIALSGYSWSRV 660
Db 601 RAPQRMAGMFQVTPAERSLEDEINRTTABDLPIFATSYIVIFLYISIALSGYSWSRV 660
QY 661 MVDKATILGCGVAVVLGAVMAAGFPSVLTGRSSLVILQVVPVLVSVGADNIFIFVLE 720
Db 661 MVDKATILGCGVAVVLGAVMAAGFPSVLTGRSSLVILQVVPVLVSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGLALTMPAPVTRTALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGLALTMPAPVTRTALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCVKPQELPPPGQGGELLLGFPQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCVKPQELPPPGQGGELLLGFPQKAYAPPL 840
QY 841 LHWITRGVVALFLALFGVSLYSMCHISVGLDQELALPKDSVLLDLYFLFLNRYFEVGPV 900
Db 841 LHWITRGVVALFLALFGVSLYSMCHISVGLDQELALPKDSVLLDLYFLFLNRYFEVGPV 900
QY 901 YFVTTGLGNFSEAGMNAICSSAGCNFSTQKQYATEFPEQSYLAIPASSWVDDFDIW 960
Db 901 YFVTTGLGNFSEAGMNAICSSAGCNFSTQKQYATEFPEQSYLAIPASSWVDDFDIW 960
QY 961 LTPSSCCRLYISGPNKDKFCFSTVNSLNCNKMCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCFSTVNSLNCNKMCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGVLT-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDGVLTVAIILSPRLEYSGTISAHCNCLYLLDSASFMAH 1080

1054 KPLKNSQDYTEALRAARELANITADLRKVGTDPAFEVPPYTTITNVFYQYLTILPEGL 1113
1081 KPLKNSQDYTEALRAARELANITADLRKVGTDPAFEVPPYTTITNVFYQYLTILPEGL 1140
1114 FMLSCLVPTFAVSCLLGLDLRGLMLLSIVMLVDTVGFMAWDISYNAVSLINLVS 1173
1141 FMLSCLVPTFAVSCLLGLDLRGLMLLSIVMLVDTVGFMAWDISYNAVSLINLVS 1200
1174 AVGNSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1233
1201 AVGNSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1260
1234 LIQIIEFFRLNLIITLGLLHGLVFLPVLISYVGPDPNPALALEQKRAEEAAVAVWVASC 1293
1261 LIQIIEFFRLNLIITLGLLHGLVFLPVLISYVGPDPNPALALEQKRAEEAAVAVWVASC 1320
1294 NHPRSVSTADNIYVNHSPSGSIKGAISNPLPNNGRQF 1332
1321 NHPRSVSTADNIYVNHSPSGSIKGAISNPLPNNGRQF 1359

RESULT 7
ADJ27337 standard; protein; 1359 AA.
XX
AC ADJ27337;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human NPC1L1.
XX
KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SBD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.
XX
OS Homo sapiens.
XX
XX WO2004009772-A2.
XX
XX 29-JAN-2004.
XX
XX 17-JUL-2003; 2003WO-US022467.
XX
XX 19-JUL-2002; 2002US-0397442P.
XX
XX (SCHE) SCHERING CORP.
XX
XX Altmann SW, Murgolo NJ, Wang LO, Graziano MP;
XX
XX WPI; 2004-132945/13.
XX
XX N-PSDB; ADJ27336.
XX
XX GENBANK; AF192522.
XX
XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
XX or identifying antagonists of NPC1L1 for inhibiting intestinal
XX cholesterol absorption in a subject, or for treating elevated serum
XX cholesterol or stroke.
XX
XX Example 7; SEQ ID NO 44; 125pp; English.
XX
XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1
XX (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
XX a motif ADJ27331 which acts as a trans-golgi network to plasma membrane
XX transport signal, and which exhibits limited tissue distribution and
XX gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
XX sterol regulated element binding protein 1 (SREBP1) binding consensus
XX sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
XX sensing domain (SBD) which is involved in sensing cholesterol levels,
XX possibly by a mechanism which involves direct cholesterol bonding. NPC1L1

CC has 42% amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.
XX
SQ Sequence 1359 AA;
Query Match 99.5%; Score 6872.5; DB 8; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 MAEAGLRCWLLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPGLSGLMTLSNVCSLSN 60
DB 1 MAEAGLRCWLLWALLRLAQSEPYTTIHQPGYCAFVDECGKNPGLSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPLRYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPLRYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTCMGYGSALCNAQRWLNFGDGTGNGLAPLDITTHLLEPGQAVSGIOPLNEGV 240
DB 181 ATLAVGTCMGYGSALCNAQRWLNFGDGTGNGLAPLDITTHLLEPGQAVSGIOPLNEGV 240
QY 241 ARCNEQGDVATCSCQDCAASCPAIAIPQALDSTFYLGQMPGSLVLIILICSVFVAVVTI 300
DB 241 ARCNEQGDVATCSCQDCAASCPAIAIPQALDSTFYLGQMPGSLVLIILICSVFVAVVTI 300
QY 301 LLVGFRAVAPARDKSMVDPKKGTSLSKLSSTHTILGQFFQCGMTWASWPLTILVLSV 360
DB 301 LLVGFRAVAPARDKSMVDPKKGTSLSKLSSTHTILGQFFQCGMTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVPTTELTTPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVPTTELTTPVELWSAPNSQARSEKAFHQHFGPPFRTNQVILTAPNRSY 420
QY 421 RYDSILLGPKNFGSILDLLELLELQERLRHLQVWSPEAQRNISLQDI CYAPLNPDNT 480
DB 421 RYDSILLGPKNFGSILDLLELLELQERLRHLQVWSPEAQRNISLQDI CYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDMKDHPLYCANAPLTPKDGTTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDMKDHPLYCANAPLTPKDGTTALAL 540
QY 541 SCWADYGAPVPPFLAIGYKGYDSEAEALIMTFSINNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCWADYGAPVPPFLAIGYKGYDSEAEALIMTFSINNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFWTFTFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISALAGSYSSWSKV 660
DB 601 RAFQRMAGMFWTFTFAERSLEDEINRTTAEDLPITPATSYIIVIFLYISALAGSYSSWSKV 660
QY 661 MVDKATLGLGVAVVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSYGADNIFIFVLE 720
DB 661 MVDKATLGLGVAVVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSYGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHTRALGRVAPSMMLCSLSEATCFFLGLATPMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVIHTRALGRVAPSMMLCSLSEATCFFLGLATPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKVQPELPPPGQEGLLGFFOKAVAPFL 840
DB 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKVQPELPPPGQEGLLGFFOKAVAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTLTGYNFSSEAGNNAICSSAGCNFSTQKIQVATEFPQSYLAIPASSWVDDFDW 960
 DB 901 YFVTLTGYNFSSEAGNNAICSSAGCNFSTQKIQVATEFPQSYLAIPASSWVDDFDW 960
 QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCKNCMSITWGSVPSPVEQHKYLPWFINDRP 1020
 DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLCKNCMSITWGSVPSPVEQHKYLPWFINDRP 1020
 QY 1021 NIKPKGGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
 DB 1021 NIKPKGGLAAYSTSVNLTSDQVLTVAILSPRLEYSTISAHCNVYLLDSASRFMAYH 1080
 QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPVPTITNVFYEQYLTLPBGL 1113
 DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPVPTITNVFYEQYLTLPBGL 1140
 QY 1114 FMLSCLVPTFAVSCLLGLDLSRLSGLNLLSVMLVDTVGPMAWDSYNVAVSLINLVS 1173
 DB 1141 FMLSCLVPTFAVSCLLGLDLSRLSGLNLLSVMLVDTVGPMAWDSYNVAVSLINLVS 1200
 QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
 DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
 QY 1234 LIQIFFFRNLITLLGLLHGLVFLPVILSVYVGPDPNPALALEQKRAEEAAVAVVASC 1293
 DB 1261 LIQIFFFRNLITLLGLLHGLVFLPVILSVYVGPDPNPALALEQKRAEEAAVAVVASC 1320
 QY 1294 NHPRSVSTADNLYVNHSPSGSIKGAISNFLPNNGRQF 1332
 DB 1321 NHPRSVSTADNLYVNHSPSGSIKGAISNFLPNNGRQF 1359

RESULT 8

ADY60780
 ID ADY60780 standard; protein; 1359 AA.
 XX
 AC ADY60780;
 XX
 XX 19-MAY-2005 (first entry)
 DT
 XX Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 44.
 DE
 XX Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
 KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
 KW metabolic disorder; atherosclerosis; cardiovascular disease;
 KW coronary artery disease; cerebrovascular ischemia; neurological disease;
 KW arteriosclerosis; transgenic animal.
 XX
 OS Homo sapiens.
 XX
 XX WO2005015988-A1.
 PN
 XX 24-FEB-2005.
 PD
 XX 16-DEC-2003; 2003WO-US040113.
 PF
 XX 17-JUL-2003; 2003US-00621758.
 PR
 XX 22-AUG-2003; 2003US-00646301.
 PR
 XX 16-SEP-2003; 2003US-00663208.
 XX
 XX (SCHE) SCHERING CORP.
 XX
 PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
 XX WPI; 2005-284403/29.
 DR N-PSDB; ADY60779.
 XX
 XX Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
 PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
 PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
 XX

Disclosure; SEQ ID NO 44; 146pp; English.

PS The invention relates to an isolated Niemann-Pick C1-like protein 1
 XX (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
 CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (rat,
 CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
 CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
 CC recombinant vector comprising the polynucleotide, a host cell comprising
 CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
 CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
 CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
 CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
 CC produce any functional NPC1L1 protein), an offspring or progeny of the
 CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
 CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
 CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject,
 CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
 CC dosage form, and information indicating that NPC1L1 is a target of
 CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
 CC sterol or Salpha-sterol absorption in a subject (involving reducing the
 CC level of expression of NPC1L1 in the subject), identifying an antagonist
 CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
 CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
 CC of NPC1L1 which is useful for inhibiting or decreasing the level of
 CC NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
 CC mouse, rat or human. The level of expression of NPC1L1 in the subject is
 CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
 CC useful for screening a sample for intestinal sterol or Salpha-sterol
 CC absorption antagonist. The NPC1L1 antagonists identified are useful for
 CC treating medical conditions caused or mediated by NPC1L1, e.g.,
 CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
 CC arteriosclerosis. The present sequence represents a mammalian NPC1L1
 CC protein.
 XX

SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLLLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
 DB 1 MAEAGLRGWLWALLLLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
 QY 61 TPARKITGDHLILLOKICPRLYTCGNTQACCSAKOLVSLASLSITKALLTRCPACSNF 120
 DB 61 TPARKITGDHLILLOKICPRLYTCGNTQACCSAKOLVSLASLSITKALLTRCPACSNF 120
 QY 121 VNLHCHNTCSPNQSLFINVTVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
 DB 121 VNLHCHNTCSPNQSLFINVTVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
 QY 181 ATLAVGTWCGYVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLRPGQAVGSGIQPLNEGV 240
 DB 181 ATLAVGTWCGYVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLRPGQAVGSGIQPLNEGV 240
 QY 241 ARCNEQGDVATCSQDDCAASCAPAIARPOALDSTFYLGQMPGSLVLIIILCSVPVAVTI 300
 DB 241 ARCNEQGDVATCSQDDCAASCAPAIARPOALDSTFYLGQMPGSLVLIIILCSVPVAVTI 300
 QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGWGTWVASWPLTILVLSV 360
 DB 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQPFQGWGTWVASWPLTILVLSV 360
 QY 361 IPVVALAAGLVFTTELTTDPVELWSAPNSQARSEKAFHDQHPGPPFRNQVILTAPNRSSY 420
 DB 361 IPVVALAAGLVFTTELTTDPVELWSAPNSQARSEKAFHDQHPGPPFRNQVILTAPNRSSY 420
 QY 421 RYDSLLGPKNFSGILDLDLLELLELQERLHIVQWSPRAQRNISLQDICYAPLNPDNT 480
 DB 421 RYDSLLGPKNFSGILDLDLLELLELQERLHIVQWSPRAQRNISLQDICYAPLNPDNT 480
 QY 481 SLYDCCINSLIQYFQNNRTLLILLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDTALAL 540

Db 481 SLIDCCINSLLQYQNNRTLLLTANQTLAQTSQVQDKHFLYCANAPLTFKDGTTALAL 540
Qy 541 SCMDYGAAPVPPFLAIGGYKGYKDYSEBALIMTSINNYNYPAGDPRLAQAKLWEAFLEEM 600
Db 541 SCMDYGAAPVPPFLAIGGYKGYKDYSEBALIMTSINNYNYPAGDPRLAQAKLWEAFLEEM 600
Qy 601 RAFORWAGMFQVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSYSSWSRV 660
Db 601 RAFORWAGMFQVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGSYSSWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVPPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVPPFLVLSVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGALGRVAPSMLLCSLSEAI CFFLGCALTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGALGRVAPSMLLCSLSEAI CFFLGCALTPMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKVQELPPPGQCGEGLLGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCCKVQELPPPGQCGEGLLGFFQKAYAPPL 840
Qy 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYEFGAPV 900
Db 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYEFGAPV 900
Qy 901 YFVTLTGYNFSSEAGNNAICSSAGCNFSFTOKIOYATEFFPROSYLAIPASSWVDDFDW 960
Db 901 YFVTLTGYNFSSEAGNNAICSSAGCNFSFTOKIOYATEFFPROSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQFHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQFHKYLPMFLNDRP 1020
Qy 1021 NIKCPKGLAAVSTSVNLTSDDQVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAVSTSVNLTSDDQVLDTVAILSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPDPAFEVPPYTTTNVFEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPDPAFEVPPYTTTNVFEQVLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWIDSYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWIDSYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKENTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKENTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFFRMLNLLTLGLLHGLVPLVILSYVGFDPVNPALALQKRAEAAVAVMVASCP 1293
Db 1261 LIQIFFRMLNLLTLGLLHGLVPLVILSYVGFDPVNPALALQKRAEAAVAVMVASCP 1320
Qy 1294 NHPSRVSTADNIYVNHSEFSGSIKAGAISNPLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSEFSGSIKAGAISNPLPNNGRQF 1359

RESULT 9
AEB93610
ID AEB93610 standard; protein; 1359 AA.

XX AC AEB93610;

DT 06-OCT-2005 (first entry)

XX DE Human NPC1L1 protein.

XX KW Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
KW arteriosclerosis; db.

XX OS Homo sapiens.
XX WO2005069900-A2.
XX 04-AUG-2005.
XX 14-JAN-2005; 2005WO-US001469.
XX 16-JAN-2004; 2004US-0537341P.
XX (MERI) MERCK & CO INC.
XX Garcia-Calvo M;
XX WPI; 2005-564070/57.
XX N-PSDB; AEB93609.
XX GENBANK; AF192522.
XX Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by
XX contacting NPC1L1 with detectably labeled substituted 2-azetidinone
XX glucuronide and a candidate compound and determining if compound binds to
XX human NPC1L1.
XX Disclosure; SEQ ID NO 44; 215pp; English.
XX The invention relates to identifying a ligand of NPC1L1. The method
XX involves contacting human NPC1L1 with a detectably labeled substituted 2-
XX azetidinone glucuronide and a candidate compound and determining whether
XX the candidate compound binds to human NPC1L1. In identifying a ligand of
XX NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
XX nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
XX is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands
XX are useful for stimulating or blocking the activity of NPC1L1, and for
XX treating conditions caused or mediated by NPC1L1. It is useful for
XX reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
XX disease, stroke, or arteriosclerosis. The present sequence represents the
XX amino acid sequence of human NPC1L1 as disclosed under GenBank accession
XX number AF192522.
XX Sequence 1359 AA;
XX Query Match 99.5%; Score 6872.5; DB 9; Length 1359;
XX Best Local Similarity 97.9%; Pred. No. 0;
XX Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCFAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCFAFYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKOLVSEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKOLVSEASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCINTCPNQSLFINVTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCINTCPNQSLFINVTRVAQLGAGOLPAVAYEAPYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGGDDVATCSQDCQAACPAIARPAQALDSTFYLQMPGSLVLIILCSVFVAVTI 300
Db 241 ARCNEQGGDDVATCSQDCQAACPAIARPAQALDSTFYLQMPGSLVLIILCSVFVAVTI 300
Qy 301 LLVGFRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGFRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTTDTPVELWSAPNSQARSEKAFHDQHFQFFRTNQTIVLTAPNRSSY 420
Db 361 IPVVALAAGLVFTLTTDTPVELWSAPNSQARSEKAFHDQHFQFFRTNQTIVLTAPNRSSY 420

```
QY 421 RYDSLLGPKNFGSGLDLDLLELELEQLERLRLHQLVMSPEAQRNISLQDICVAPLNPDNT 480
D 421 RYDSLLGPKNFGSGLDLDLLELELEQLERLRLHQLVMSPEAQRNISLQDICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLALTANOTLMGOTSQVDWKDHFLYCANAPLTKDGTALAL 540
D 481 SLYDCCINSLLQYFQNNRTLLALTANOTLMGOTSQVDWKDHFLYCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFFFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAALWEAFLEEM 600
D 541 SCWADYGAPVFFFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAALWEAFLEEM 600
QY 601 RAFORMAGNFQVTFPAERSLDEINRTTAEDLPFATSYIYIYISIALGYSYSWSRV 660
D 601 RAFORMAGNFQVTFPAERSLDEINRTTAEDLPFATSYIYIYISIALGYSYSWSRV 660
QY 661 MYDSKATLGLGCVAVVLGAVMAAMGFPSYLGTRSSILVLOVPPFLVLSVGADNIFIVLE 720
D 661 MYDSKATLGLGCVAVVLGAVMAAMGFPSYLGTRSSILVLOVPPFLVLSVGADNIFIVLE 720
QY 721 YQRLPRRPGEPREHVIGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAPVTRTALTSGLAV 780
D 721 YQRLPRRPGEPREHVIGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAPVTRTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQEAERLDVCCCKPQELPPPPQGGELLLGFPKAYAPPL 840
D 781 ILDFLLQMSAFVALLSLDSKRQEAERLDVCCCKPQELPPPPQGGELLLGFPKAYAPPL 840
QY 841 LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
D 841 LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATFEPQSYLAIPASSWVDVDFIDW 960
D 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSFTQKIYATFEPQSYLAIPASSWVDVDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLPFLNDRP 1020
D 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLPFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSQGVLT-----ASRPMAYH 1053
D 1021 NIKCPKGLAAYSTSVNLTSQGVLT-----ASRPMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITTNVFEQVLTILPGL 1113
D 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITTNVFEQVLTILPGL 1113
QY 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITTNVFEQVLTILPGL 1140
D 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITTNVFEQVLTILPGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSYVILVDTVGMALWDISYNAVSLINLVS 1173
D 1114 FMLSCLVPTFAVSCLLGLDLRLSGLLMLLSYVILVDTVGMALWDISYNAVSLINLVS 1173
QY 1174 AVGMSVEFVSHITRSPFAISTKTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
D 1174 AVGMSVEFVSHITRSPFAISTKTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
QY 1201 AVGMSVEFVSHITRSPFAISTKTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
D 1201 AVGMSVEFVSHITRSPFAISTKTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
QY 1234 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1293
D 1234 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1293
QY 1261 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1320
D 1261 LIQIFEFRLNLTILGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEAAVAVMVASCP 1320
QY 1294 NHPSRVSTADNLYVNHSPSGSKGAGAI SNFLPNNGRQF 1332
D 1321 NHPSRVSTADNLYVNHSPSGSKGAGAI SNFLPNNGRQF 1359
```

RESULT 10

ID ABG22693

XX standard; protein; 1344 AA.

AC ABG22693;

XX 18-FEB-2002 (first entry)

```
XX Novel human diagnostic protein #22684.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
OS Homo sapiens.
PN WO200175067-A2.
PD 11-OCT-2001.
PF 30-MAR-2001; 2001WO-US008631.
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS86880.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 53052; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1344 AA;
Query Match 94.6%; Score 6536; DB 4; Length 1344;
Best Local Similarity 96.7%; Pred No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;
QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
D 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
D 61 TPARKITGDHLLILLOKICPRLTYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNPQSLFINVTRVAQGLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 180
D 121 VNLCHNTCSNPQSLFINVTRVAQGLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAVGTMCVGYGSALCNQARWLPFGDGTGNGLAPLDTTFHLEPGQVGSQIPLNEGV 240
D 181 ATLAVGTMCVGYGSALCNQARWLPFGDGTGNGLAPLDTTFHLEPGQVGSQIPLNEGV 240
```

Db 181 ATLAVGTCVGYGALCAQRLWLFQDGTGNGLAPLDITFHLLEPGQAVGSGIOPLNEGV 240

Qy 241 ARCHESQDDVATCSCQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILICSVFVAV 298

Db 241 ARCHESQDDVATCSCQDCAASCPXPAPRSTPSTW--RQMPGSLVLIILICSVFVAV 298

Qy 299 TILLVGFVAPARDKSKWDPKKTSLSDKLSFSTHTLLGQFQGWGTWVASWPLTILVL 358

Db 299 TILLVGFVAPARDKSKWDPKKTSLSDKLSFSTHTLLGQFQGWGTWVASWPLTILVL 358

Qy 359 SVIPVVALAAGIVFTLTTDPVELWSAPNSQARSEKAFHQHFGFPFFFTNQVILTAPNRS 418

Db 359 SVIPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHQHFGFPFFFTNQVILTAPNRS 418

Qy 419 SYRYDSLLGPKNFSGIILDLILLLELQERLHLQWSPEAQNISLQIDICVAPLNPD 478

Db 419 SYRYDSLLGPKNFSGIILDLILLLELQERLHLQWSPEAQNISLQIDICVAPLNPD 478

Qy 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTL 538

Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTL 538

Qy 539 ALSCHADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 598

Db 539 ALSCHADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 598

Qy 599 EMRAFORMMAGMFQVFTTATERSLEDEINRTTAEDLPFATSYVIVIFLYISLALGYSYWS 658

Db 599 EMRAFORMMAGMFQVFTTATERSLEDEINRTTAEDLPFATSYVIVIFLYISLALGYSYWS 658

Qy 659 RVWVDSKATILGCGVAVNTGAVWAMGPFYSGIRSSVILQVPLVLSVGADNIFIV 718

Db 659 RVWVDSKATILGCGVAVNTGAVWAMGPFYSGIRSSVILQVPLVLSVGADNIFIV 718

Qy 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSL 778

Db 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEAICTFFLGALTPMPAVRTFALTSL 778

Qy 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVQBELPPGQBGELLGFFQKAYAP 838

Db 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVQBELPPGQBGELLGFFQKAYAP 838

Qy 839 FLHLWITRGVW----LLLFLALPGVSLYSVGLDQELALPKDSYLLDYFLFLNRYF 894

Db 839 FLHLWITRGVWVPSQLLLFLALPGVSLYSVGLDQELALPKDSYLLDYFLFLNRYF 898

Qy 895 EVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNNFSTQKIQVATEPPEQSYLAIPASSWV 954

Db 895 EVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNNFSTQKIQVATEPPEQSYLSLPASSWV 958

Qy 955 DDFIDMLTPSSCCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRSPVEQFHXYLPW 1014

Db 955 DDFIDMLTPSSCCRLYISGNKDKFCPSTVNSLNCNKMSITMGSVRSPVEQFHXYLPW 1018

Qy 1015 FLNDRPNIKCPKGLAAYSTVNLITSDQVLAERFMAYHKPLKNSQDYTEALRAARELAA 1074

Db 1015 FLNDRPNIKCPKGLAAYSTVNLITSDQVLAERFMAYHKPLKNSQDYTEALRAARELAA 1078

Qy 1075 NITADLRKVPGTDPAPFVFFYTTITNVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLD 1134

Db 1075 NITADLRKVPGTDPAPFVFFYTTITNVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLD 1138

Qy 1135 LRSGLLNLLSIWMLVDTVGFMALWDISYNVAVSLINLVSAGMVSFVSHITSPALSTK 1194

Db 1135 LRSGLLNLLSIWMLVDTVGFMALWGISYNVAVSLINLVSAGMVSFVSHITSPALSTK 1198

Qy 1195 PTWLERAKEATISMGSAVFAGVAVMTNLPGLIIVLGLAKAQIQLIFFFLNLIITLLGLHG 1254

Db 1195 PTWLERAKEATISMGSAVFAGVAVMTNLPGLIIVLGLAKAQIQLIFFFLNLIITLLGLHG 1258

Qy 1255 LVFLPVLISYVGPDPVNPALAEQKRAEA-----VAAMVWASCNHPSRVST 1301

Db 1255 LVFLPVLISYVGPDPVNPALAEQKRAEBEGGSGHGPLAQITPPESPOLTTSMT 1313

RESULT 11

ADJ27295

ID ADJ27295 standard; protein; 1331 AA.

XX

AC ADJ27295;

XX

DT 20-MAY-2004 (first entry)

XX

DE Rat NPC1L1.

XX

KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promotor;

KW sterol regulated element binding protein 1; SREBP1;

KW binding consensus sequence; transmembrane domain; sterol-sensing domain; SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal; cholesterol absorption; serum cholesterol; hyperlipidaemia; atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX

OS Rattus sp.

XX

FN WO2004009772-A2.

XX

PD 29-JAN-2004.

XX

PF 17-JUL-2003; 2003WO-US022467.

XX

PR 19-JUL-2002; 2002US-0397442P.

XX

PA (SCHE) SCHERING CORP.

XX

PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX

DR WPI; 2004-132945/13.

DR N-PSDB; ADJ27294.

XX

PT New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting or identifying antagonists of NPC1L1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

XX

PS Claim 1; SEQ ID NO 2; 125pp; English.

XX

CC This sequence represents a Niemann-Pick disease, type C1, gene-like 1 (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and gastrointestinal abundance. The human NPC1L1 promoter sequence contains a sterol regulated element binding protein 1 (SREBP1) binding consensus sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-sensing domain (SSD) which is involved in sensing cholesterol levels, possibly by a mechanism which involves direct cholesterol binding. NPC1L1 has 42% amino acid sequence homology to human NPC1, a receptor responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or polynucleotides are useful for detecting or identifying antagonists of NPC1L1, which can be used for inhibiting intestinal cholesterol absorption in a subject, or for treating medical conditions including elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary heart disease, stroke or arteriosclerosis.

XX

SQ Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 8; Length 1331;

Best Local Similarity 77.8%; Pred. No. 0;

Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGLRGMLLWALLRLAQSEPYTHPGYCAFYDECGKPELSGLTSLNVSCLSN 60

Db 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKPELSGLTSLNVSCLSN 59

Qy 61 TPARKITGHLILLOKICPRLVTPGNTO-ACCSAKQLVSLASLSITKALLTRCPACSDN 119

Db 60 TPAHVHTGEHLAIIQRICPLRYNGPNTTFACCKTKQLLSLESSMSITKALLTRCPACSDN 119
QY 120 FVNHLCHNTCSNPQSLFINVTRVAOLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRVRVPA 179
Db 120 FVSLHCHNTCSNPQSLFINVTRVVERGAGEPPAVVAEYAFYQHSFAEKAYESCSQVRIPA 179
QY 180 AATLAVGTMCVYGSALCNAQRMWLNFGDGTGNGLAELDITFHLLEPGQAVSGIGIQLNEG 239
Db 180 AASLAVGSMCGVYGSALCNAQRMWLNFGDGTGNGLAELDITFHLLEPGQALPDGIQIPLNGK 239
QY 240 VARNCSQGDVATCSQDCAACPAIARPOALDSTFYLGOMPGSLVLIILCSVPAVVT 299
Db 240 IAPCNESQGDSDAVSCQDCAACPVIPPEALRPSFTYMGFMGWLALIIIFTAVFVLLS 299
QY 300 ILAVFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWTWASMLPTIIVLS 359
Db 300 AVLVRVSVNRNKNKAEGPOEAPKLPKHKLSPHITLGRFFQNGWTRVASWPLTVLALS 359
QY 360 VIPVVALAAGLVTELTDDVELWSAPNSQARSEKAFHDQHFQPPFRTRNQVILTPANRSS 419
Db 360 FIVVIALAAGLTIFELTDDVELWSAPNSQARSEKSFHDEHFGFPFRTRNQIFVTARNRSS 419
QY 420 YRYDSLGLGPKFSGILDLLLELLELOERLRLHLOWSPQAORNI-SLODICVAPLNPDN 479
Db 420 YKDSLGLGPKFSGILDLLLELLELOERLRLHLOWSPQAORNI-SLODICVAPLNPDN 479
QY 480 TSLYDCINSLLOYFQNNRTLLLTANQITLMGTQSOVDKMDHFLYCANAPLTPKQGTALA 539
Db 480 TSLSDCCVNSLLQYFQNNRTLLMTANQITLMGTSLVDKMDHFLYCANAPLTPKQGTSLA 539
QY 540 LSCMADYGAPVPFPLAIGYKGDYSEAEALIMTFSNNYPAGDPRLAQAUKWEAFLEB 599
Db 540 LSCMADYGAPVPFPLAVGGYQGYDYSEAEALITFSLNNYPADDPDMAQAUKWEAFLEB 599
QY 600 MRAFORMAGMFOVTFATRESLEDEINRTTAEPLIPATSYIVIVLYISLALGSSYSMSR 659
Db 600 MESFORNTSKFOVAFSAKRESLEDEINRTTIQDLPVFAVSIIIVLYISLALGSSYSRCSR 659
QY 660 VMVDSKATILGAGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIIFVL 719
Db 660 VAVESKATILGAGVAVVLGAVLAAMGFYSLGVPSLVIQVVPFLVLAAGADNIIFVL 719
QY 720 EYQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAI-CFFLGLATMPAVRTPALTSGLA 779
Db 720 EYQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAI-CFFLGLATMPAVRTPALTSGLA 779
QY 780 VILDFFLQMSAFVALLSLDSKQEAERLDVCCCKVQBELPPQCGGLLLGFPKAYAPF 839
Db 780 IILDFFLQMTAFVALLSLDSKQEAERLDVCCCKVQBELPPQCGGLLLGFPKAYAPF 839
QY 840 LLHWITRGVVLILFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPRVGGAP 899
Db 840 LLHRFTRPVVMLLFULFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGPP 899
QY 900 VYFVTTLYGNFSEAGMNAICSSAGCNFSPFTQKIQYATEFPEQSYLAIPASSWVDDFD 959
Db 900 VYFVTTLYGNFSEAGMNAICSSAGCNFSPFTQKIQYATEFPEQSYLAIPASSWVDDFD 959
QY 960 WLTP-SSCCRLYISGNKDKFCBSTVNSLNCNKMSITWGSVRPSVEGFHYKLPFLND 1018
Db 960 WLTPSSCCRLYIRGPHDDEFCEPSTDTSFNCLNCKMNRITLGPVRPTAEQFHYKLPFLND 1019
QY 1019 RPNKCPKGLAAYSTSVNLTSQGLASRFMAYHKPLKNSODYTEALRAARELANITA 1078
Db 1020 PNNIRCPKGLAAYRSTSVNLSGQVIAQFMAHYKPLKNSQDFTBALRASULLANIYA 1079
QY 1079 DLKRVGTPDAPFVFPYITINVFYEQYLTLPLEGLPMLSLCLVPTFAVSCLLGLDLRSG 1138
Db 1080 DLKRVGTPDNPVFPYITINVFYEQYLTLPLEGLPMLSLCLVPTFAVSCLLGLDMCSG 1139
QY 1139 LNLLSIVMLVDVTGFMALWISYNAVSLINLVSAGVMSVFPVSHITRSPAISTKPTWL 1198
Db 1140 ILNLLSIIMLVDITGLMAVWGISYNAVSLINLVTAVGMSVFPVSHITRSPAISTKPTRL 1199

QY 1199 ERAKEATISMGSVAFAGVAMTNLPGILVLGLAKAQLTIQIFFFRNLNLLITLLGLHLGVFL 1258
Db 1200 ERAKDATVFMGSVAFAGVAMTNPPGILVLGLFAQAQLTIQIFFFRNLNLLITLLGLHLGVFL 1259
QY 1259 PVLSLVGDPDVPALALBOKRABEAAVAAVWVASCPNHPSRVSTADNIYVNHSEFSI-KG 1317
Db 1260 PVLSLVGDPDVPALALBOKRABEAAVAAVWVASCPNHPSRVSTADNIYVNHSEFSI-KG 1316
QY 1318 AGAISNPLPNNGROF 1332
Db 1317 ANAARSSLPKSDORF 1331

RESULT 12

ADY60738
ID ADY60738 standard; protein; 1331 AA.
XX
AC ADY60738;
XX
DT 19-MAY-2005 (first entry)
XX
DE Rat Niemann-Pick C1-like protein 1 (NPC1L1) protein.
XX
KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX
OS Rattus sp.
XX
PN WO2005015988-A1.
XX
PD 24-FEB-2005.
XX
PF 16-DEC-2003; 2003WO-US040113.
XX
PR 17-JUL-2003; 2003US-00621758.
PR 22-AUG-2003; 2003US-00646301.
PR 16-SEP-2003; 2003US-00663208.
XX
PA (SCHE) SCHERING CORP.
XX
PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX
DR WPI; 2005-284403/29.
XX
DR N-PSDB; ADY60737, ADY60746.
PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
XX
PS Claim 30; SEQ ID NO 2; 146pp; English.
XX
CC The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the
CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of
CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
CC sterol or Salpha-sterol absorption in a subject (involving reducing the

CC level of expression of NPC1l1 in the subject), identifying an antagonist
 CC of NPC1l1 and an isolated mammalian cell which lacks a gene encoding a
 CC functional NPC1l1 protein. NPC1l1 is useful for identifying an antagonist
 CC of NPC1l1 which is useful for inhibiting or decreasing the level of
 CC NPC1l1 mediated sterol or Salpha-stanol uptake in a subject such as
 CC mouse, rat or human. The level of expression of NPC1l1 in the subject is
 CC reduced by mutating NPC1l1 in the subject. The NPC1l1 knockout mouse is
 CC useful for screening a sample for intestinal sterol or Salpha-stanol
 CC absorption antagonist. The NPC1l1 antagonists identified are useful for
 CC treating medical conditions caused or mediated by NPC1l1, e.g.,
 CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
 CC arteriosclerosis. The present sequence represents a mammalian NPC1l1
 CC protein.
 XX
 SQ Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;
 Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSPPYTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
 DB 1 MAEAWL-GWLLWALLSAAQGLYTPKHEAGVCTFYECGKNPELSGSLTSLSNVSCLSN 59
 QY 61 TPARKITGDHLILQKICPLRYTGNTQ-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
 DB 60 TPARHVTGBHALLORICPRLYNGENTTPACCSTKQLLSLESSMSITKALLTRCPACSDN 119
 QY 120 FYNLHCNHTCSNQSLFINVTRVAQGLGALPAVAYEAFYQHSFAEQSYDCSCSRVRPVA 179
 DB 120 FVSLHCNHTCSQDLSFINVTRVERGAGEPPAVYAYEAFYQHSFAEKAYESCQRIPA 179
 QY 180 AATLAVGTWCYVGSALCNAQRLNPGQDGTGNGLAPLDTITFHLBPQAVGSGIQLNEG 239
 DB 180 AASLAVGSCYVGSALCNAQRLNPGQDGTGNGLAPLDTITFHLBPQALPDGIQLNGK 239
 QY 240 VARCHESQDDVATCSQDCAPAIARPOALDSTFYLGQMPGSLVLIILCSVPVAVT 299
 DB 240 IAPCNESQDSDSVAVCSQDCAPAIARPPPEARLPSFYNGRMPGMLALIIITFAVFLVS 299
 QY 300 ILLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQGMGTWVASWPLTILVLS 359
 DB 300 AVLRLRVVNNKNAEGPQAPKLPKHKLSPHTILGRFFQNGWTRVASWPLTILVLS 359
 QY 360 VIPVVALAAGLVTELTITDPVELWSPNSQARSEKAFHQHFGCPFRPTNQVILTAPNRSS 419
 DB 360 FIVIALAAGLTFIELTITDPVELWSPKQARSEKSFHDEHFGFPFRPTNQIFVTARNSS 419
 QY 420 YRYDSLGLGKPFSGILDLDLELELOERLRLQVWSPQARNISLODICVAPLNPDN 479
 DB 420 YKYDSLGLGKPFSGILSDLELELOERLRLQVWSPQARNISLODICVAPLNPDN 479
 QY 480 TSLYDCCINSLLQYFQNNRTLLMTANQTLNGQTSVDMKHFLYCANAPLTFKQGTALA 539
 DB 480 TSLSDCCVNSLLQYFQNNRTLLMTANQTLNGQTSVDMKHFLYCANAPLTFKQGTSLA 539
 QY 540 LSCWADYGAPVFPFLAIGVKGKDYGEAEALIMTFSINNYPADDPRLAQAKLWEEAFLEE 599
 DB 540 LSCWADYGAPVFPFLAVGGYQGYDYGEAEALITFSINNYPADDPRLAQAKLWEEAFLEE 599
 QY 600 MRAFORRMAGMFOVTFEASLSDEINRTTAEOLPIFATSYIVFLYISALGYSYSWSR 659
 DB 600 MESFORNTSKFQVAFSAERSLSDEINRTTIQDLPVFAVSIIYVFLYISALGYSYSRCSR 659
 QY 660 VMYDSKATLGLGVAVVGLGAVMAAMGFFSYIGIRSSILVILQVVPFLVSLVGADNIFIVL 719
 DB 660 VAVESKATLGLGVIVVGLGAVLAAMGFYSYLGVPSSILVILQVVPFLVAVGADNIFIVL 719
 QY 720 EYQRLPRRPOEPRVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTPALTSGLA 779
 DB 720 EYQRLPRMPGEQREAHIGRTLGSVAPSMMLCSLSEALCFPLGALTMPAVRTPALTSGLA 779
 QY 780 VILDFLQMSAFVALLSLDSKRQASRLDVCCCKPQELPPPGQEGLLIGFPQKAYAPP 839

DB 780 IILDFLQMTAFVALLSLDSKRQASRPDLVLCFSTRKLPPEKKGGLLFFRKIYAPP 839
 QY 840 LLHWITGVVLLLFALFGVSLYSKMSHISVGLDDELALPKDSYLLDYDFLFLNRYPEVCAP 899
 DB 840 LLHRFIRPVVMLLFTLFGANLYLNCINVLQDELALPKDSYLLDYDFLFLNRYLEVGPP 899
 QY 900 VYFVTTLYGNFSSAGNAICSSAGCNFFSTQKIQVATPEPQSYLAIPASSWVDDFID 959
 DB 900 VYFVTTGFGNFSSAGMNATCSSAGCKSFSUTQKIQVASEFPDQSYVAIAASSWVDDFID 959
 QY 960 WLTP-SSCCRLYISGPNKDKFCPSTVNSLNCIKNCMSITMGSVRPSVQFHKYLPWFLND 1018
 DB 960 WLTPSSCCRLYIRGPHKDEFCPSTDTSFNCLKNCNRTLGVRPTAEQFHKYLPWFLND 1019
 QY 1019 RPNIKCPKGGLAAYSTSVNLSDQVLSRPMAYHKPLKNSQDYTEALRAARELAANITA 1078
 DB 1020 PENIRCPKGGLAAYRTSVNLSDQVLSRPMAYHKPLKNSQDYTEALRAARELAANITA 1079
 QY 1079 DLKVPKGTDAPEFPPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
 DB 1080 DLKVPKGTDPNFEFPPYTTISNVFQQYLTIVLPEGIFTALCFVPTFVVCYLLGLDMCS 1139
 QY 1139 ILNLISIVMILVDTVGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198
 DB 1140 ILNLISIMILVDTITGLMAVMSYNAVSLINLVTAVGMSVEFVSHITRSPAVSTKPTRL 1199
 QY 1199 ERAKEATISMGSAVPAGVAMTNLPGILVGLAKAQLIQIFPRLNLLITLGLLHGLVPL 1258
 DB 1200 ERAKDATVFMGSAVPAGVAMTNFPGILTLGFAQQLIQIFPRLNLLITLGLLHGLVPL 1259
 QY 1259 PVLLSYGDPDVPNALALROKRAEAAVAVMVASCPNHPSRVSTADNIYVNHSEFESI-KG 1317
 DB 1260 PVLLSYLGPDVQALVQEEKLASEA-AVAPFSCPYFSPADADAN--VNYGFAPELANG 1316
 QY 1318 AGAISNLFNNGRQF 1332
 DB 1317 ANAARSLPKSDQKF 1331

RESULT 13
 AEB93568
 ID AEB93568 standard; protein; 1331 AA.
 XX
 AC AEB93568;
 XX 06-OCT-2005 (first entry)
 XX Rat NPC1l1 protein.
 XX Protein engineering; NPC1l1; antidiabetic; antiarteriosclerotic;
 KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
 KW arteriosclerosis; rat.
 XX Rattus sp.
 XX WO2005069900-A2.
 XX 04-AUG-2005.
 XX 14-JAN-2005; 2005WO-US001469.
 XX 16-JAN-2004; 2004US-0537341P.
 XX (MERI) MERCK & CO INC.
 XX Garcia-Calvo M;
 XX WPI: 2005-564070/57.
 XX N-PSDB; AEB93567.
 XX Identifying ligand of NPC1l1 for stimulating the activity of NPC1l1, by
 PT contacting NPC1l1 with detectably labeled substituted 2-azetidinone

glucuronide and a candidate compound and determining if compound binds to human NPC1L1.

Example 2; SEQ ID NO 2; 215pp; English.

The invention relates to identifying a ligand of NPC1L1. The method involves contacting human NPC1L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC1L1. In identifying a ligand of NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands are useful for stimulating or blocking the activity of NPC1L1, and for treating conditions caused or mediated by NPC1L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents a rat NPC1L1, a N-glycosylated protein.

Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;

Best Local Similarity 77.8%; Pred. No. 0;

Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

1 MAEAGLRGWLWALLRLAQSPPYTHIQPGVCAPYDECGKNPELGGSLMTLSNVCSLSN 60
 1 MAAAWL-GWLLWALLLSAAQGBLYTPKHEAGVCTFEYECCKNPGLSGGLTSLSNVCSLSN 59
 61 TPARKITGDHLLQKICPRLVTGNQTO-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
 60 TPARKVTGEHLALLQRIICPRLVNGPNTTFACCKTQLLSLESSMITKALLTRCPACSDN 119
 120 FVNLCHNTCSNQSLFNIVTRVAOLGAGQAPVAVYAEAFYQHSFAEOSYSCSRVRVPA 179
 120 FVSLCHNTCPDQSLFNIVTRVVERGAGEPPAVVAYEAFYQHSFAEYSCSVRIAPA 179
 180 AATLAVGTWCGYGSALCNAQRWLNFGQDTGNGLAPLDTITPHLLPFGQAVGSGIQLNEG 239
 180 AASLAVGSMCGYGSALCNAQRWLNFGQDTGNGLAPLDTITPHLLPFGQALPDGIQLNGK 239
 240 VARNESOGDDVATCSCDCAASCPAIARPOALDSTFVYLGQMPGSLVLIILCSFPAVT 299
 240 IAPCNESOGDDSAVCSODCAASCPVIPPPEALRPSFYMGRRMPGWLIIIFTAVFVLLS 299
 300 ILLVGFVRAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFQGMGTWVWASWPLTILVLS 359
 300 AVLVLRLVSVNKNKAGSQEAPKLPKHKLSPHTILGRFQWGTVRVWASWPLTILVLS 359
 360 VIVVALLAGLVFTLTDTPVELWSAPNSQARSEKAFHDQHFQPPFRTNQVILTAPNRSS 419
 360 FIVVIALAAGLTFFIELTTPVELWSAPKQARKKESFDEHFGPPFRTNQIVFTARNSS 419
 420 YRYDSLILGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 479
 420 YKYDSLILGPKNFGSILDLFLELLELLELLELLELLELLELLELLELLELLELLEL 479
 480 TSLYDCCNSLLOXFOQNRTLLLTANOTLMGQTSOVDWKDHFLYCANAPLTFKDGTLA 539
 480 TSLSDCCVNSLLOXFOQNRTLLMTANOTLNGQTSVLDWKDHFLYCANAPLTFKDGTSIA 539
 540 LSCMADYGAPVPPFLAIGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEE 599
 540 LSCMADYGAPVPPFLAVGQYQGTDYSEAEALITFSLNNYPAGDPRLAQAKLWEEAFLEE 599
 600 MRAFORRNAGHQVFTFAERSLEDEINRTTAEIDLPIFATSVIIVFLYISLALGSVSWSR 659
 600 MESFORNTSDRFQVAFSAERSLEDEINRTTQDLFVFAVSIVFLYISLALGSYSRCSR 659
 660 VVWDSKATLGLGGVAVLVGAVMAAGFPFYSIGRSSVLQVWVPLVSGADNIFIFVL 719
 660 VAVESKATLGLGGVIVLVGAVLAAGFYISLGVPSLSVLIQVWVPLVAVGADNIFIFVL 719
 720 EYQRLPRRPGPREVHIIGALGRVAPSMMLCSLSAICFFLGTALTPMPAVRTFALTSGLA 779

Db 720 EYQRLPRRPGPREVHIIGALGRVAPSMMLCSLSAICFFLGTALTPMPAVRTFALTSGLA 779
 Qy 780 VILDLLQWSAPVALLSLSDSKBOEASRLDVCCVCPQELPPPGQEGILLGFQKAYAPF 839
 Db 780 IILDLLQWTAVALLSLSDSKBOEASRLDVCCVCPQELPPPGQEGILLGFQKAYAPF 839
 Qy 840 LLIHMTIRGVVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVGAP 899
 Db 840 LLIHMTIRGVVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVGAP 899
 Qy 900 VYFVTTLGVNPSSEAGMAICSSAGCNPFSTQKIQTATEPPEQSYLAIIPASSWDDFD 959
 Db 900 VYFVTTLGVNPSSEAGMAICSSAGCNPFSTQKIQTATEPPEQSYLAIIPASSWDDFD 959
 Qy 960 WLTP--SSCCRLYISGPNKPCPSVNSINCLKNCSITMGSVRPSVVEQFHKYLPWFLND 1018
 Db 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCLNCKNRTLGPVPTAEQFHKYLPWFLND 1019
 Qy 1019 RPNIKCPKGLAAYSTVNLTSQVQLASRFMAYHKPKNSQDYTEALRAARELAANITA 1078
 Db 1020 PPNIRCPKGLAAYSTVNLSSDQVIAQFVAYHKPLNSQDFTFALRASRLAANITA 1079
 Qy 1079 DLRKVPGTDPAPFVPPYITINVFYQYLTITLPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
 Db 1080 DLRKVPGTDPFVPPYITINVFYQYLTITLPEGLFMLSCLVPTFAVSCLLGLDMCSG 1139
 Qy 1139 LILNLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWL 1198
 Db 1140 ILNLSIIMILVDTITGLMAVNGISYNAVSLINLTVAVGMSVEFVSHITRSPAISTKPTWL 1199
 Qy 1199 ERAKEATTSMGSAVPAVGMNTLPGILVLGLAKAQLIOIFPFRMLLITLGLLHGLVPL 1258
 Db 1200 ERAKATVPMGSAVPAVGMNTLPGILVLGLAKAQLIOIFPFRMLLITLGLLHGLVPL 1259
 Qy 1259 PVILSYVGPDPNPALEBQKRAEAAVAVMVASCPNHPFSRVSTADNIVYNHSEFSGI-KG 1317
 Db 1260 PVILSYLGPDPNPALEBQKRAEAAVAVMVASCPNHPFSRVSTADNIVYNHSEFSGI-KG 1317
 Qy 1318 AGATSNFLPNNGROF 1332
 Db 1317 ANAARSLPKSDQRP 1331
 RESULT 14
 ADJ27305
 ID ADJ27305 standard; protein; 1333 AA.
 XX
 AC ADJ27305;
 XX
 XX
 DT 20-MAY-2004 (first entry)
 XX
 DE Mouse NPC1L1.
 XX
 KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promoter;
 KW sterol regulated element binding protein 1; SREBP1;
 KW binding consensus sequence; transmembrane domain; sterol-sensing domain; SSB; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal; cholesterol absorption; serum cholesterol; hyperlipidaemia; atherosclerosis; coronary heart disease; stroke; arteriosclerosis.
 XX
 OS Mus sp.
 XX
 FN WO2004009772-A2.
 XX
 PD 29-JAN-2004.
 XX
 PF 17-JUL-2003; 2003WO-US022467.
 XX
 PR 19-JUL-2002; 2002US-0397442P.
 XX
 PA (SCHE) SCHERING CORP.

Tue Apr 11 10:57:55 2006

Qy	1262	LSYVGPDVNFPALALEQKRAEENVAAMVASCNHPSPSRVSTADNIYNHNSPEGS- IKGAGA	1320
Dδ	1263	LSYLGPDPVNQALVLEEKLEATEA-AWSEPSCPFFPADANTSDVYNYNGNFPEIPEINA	1321
Qy	1321	ISNFLPNNGRQF	1332
Dδ	1322	ASSSLPKSDQKF	1333

Search completed: April 7, 2006, 19:16:33
Job time : 208 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.
OM protein - protein search, using sw model
Run on: April 7, 2006, 19:16:49 ; Search time 52 Seconds
(without alignments)
2464.629 Million cell updates/sec
Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNLFPLNNGRQF 1332
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR 80: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2385	34.5	1278	2 T30188	Niemann-Pick C dis
2	1394	20.2	1055	2 T05663	hypothetical prote
3	1347	19.5	1456	2 T15961	hypothetical prote
4	1329	19.2	1170	2 S52525	probable membrane
5	646.5	9.4	1434	2 T30172	transmembrane prot
6	628.5	9.1	1442	2 T18538	patched protein -
7	600	8.7	915	2 S44797	F09G8.4 protein -
8	598.5	8.7	933	2 T25600	hypothetical prote
9	583.5	8.4	1182	2 T13952	membrane protein p
10	573.5	8.3	1003	2 T26746	hypothetical prote
11	572	8.3	1220	2 T18291	patched protein -
12	539.5	7.8	889	2 T29590	hypothetical prote
13	514	7.4	1405	2 T27969	hypothetical prote
14	473.5	6.9	1015	2 T15830	hypothetical prote
15	464.5	6.7	1299	1 S08119	membrane protein p
16	458.5	6.6	955	2 T21612	hypothetical prote
17	436	6.3	956	2 A89153	protein C24B5.3 [i
18	409.5	5.9	936	2 T26521	hypothetical prote
19	407	5.9	800	2 T26683	hypothetical prote
20	399.5	5.8	881	2 T31739	hypothetical prote
21	381	5.5	840	2 T33217	hypothetical prote
22	379	5.5	714	2 T16126	hypothetical prote
23	367	5.3	690	2 T23399	hypothetical prote
24	367	5.3	877	2 T24097	hypothetical prote
25	349	5.1	845	2 T25657	hypothetical prote
26	347.5	5.0	820	2 T32908	hypothetical prote
27	312	4.5	890	2 T22186	hypothetical prote
28	297.5	4.3	871	2 T28706	hypothetical prote
29	294	4.3	983	2 T21213	hypothetical prote

30	276.5	4.0	633	2 S44795	F09G8.3 protein -
31	259.5	3.8	413	2 S28276	hypothetical prote
32	224	3.2	1276	2 T18526	SREBP cleavage act
33	189.5	2.7	1154	2 T48829	related to SREBP c
34	187.5	2.7	932	2 T28820	hypothetical prote
35	186.5	2.7	1227	2 T20370	hypothetical prote
36	179.5	2.6	787	2 H71453	hypothetical prote
37	173	2.5	969	2 T33156	transport integral
38	170	2.5	746	2 A75018	probable ABC-type
39	154	2.2	823	2 B81282	multidrug resistant
40	149	2.2	1011	2 T07712	hypothetical prote
41	149	2.2	1051	2 AG3455	antibiotic transpo
42	148.5	2.1	746	2 H84301	acriflavin resista
43	146	2.1	1041	2 AC0423	
44	143	2.1	724	2 H69780	
45	141.5	2.0	1049	2 AP0561	

ALIGNMENTS

RESULT 1

T30188
Niemann-Pick C disease protein - mouse
N:Alternate names: NCP1 protein
C:Species: Mus musculus (house mouse)
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30188
R:Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison, Science 277, 232-235, 1997
A:Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis A:Reference number: Z20765, MUID:97362324; PMID:9211850
A:Accession: T30188
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1278 <LOF>
A:Cross-References: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:G2251241; C:Genetics:
A:Gene: Npc1
A:Map position: 18

Query Match 34.5%; Score 2385; DB 2; Length 1278;
Best Local Similarity 40.2%; Pred. No. 1.2e-158;
Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY	14	LLRLAQSEPTTIHQPGYCAFYDECGNPELSSMTLSVCLSNTPARKITGDHLL 73
DB	12	LLLLCPAQVFSQ-----SCWYGECH---IATGD---KRYNCKYSGPPKPLPKDGYDL 59
QY	74	LQKICPLRYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSNQ 133
DB	60	VQELCPGLFF-DNVSLLCDDIQQLQTLKSNLQLPLQLSRCPSCFYNMLTLFCELTCSHQ 118
QY	134	SLFINVTVAQLGAGQLPA----VVAYRAFYQHSFAQSYDSCSRVRVPAATLAVGTC 189
DB	119	SQFLNVTATEDYFDPKTPENKTNVKELEYVVCQSPANAMYNACRDVEAPSSNEKALGLLC 178
QY	190	GUYGSALCNAQRLNFGQDGTGNGLAPLDI-----TFHLLPQQAQVSGIQPLNEGVARN 244
DB	179	GRDARA-CNATNMIEMFNKNGQAPFTIIPVFSLSIL-----GMEPRNRNATKGCN 229
QY	245	ESQGDVATCSQDCNASC-----PATARQALDSTR-----YLGQMPGSLV 286
DB	230	RSVDVTGFCSCQDCSIVCGPKPQPPMPMPWRINGLDAMVYIMVMTVVAFLVFVFGALL 289
QY	287	LIILCSVFAVVTILLVGRVAPARDKSMVDPKKGTSLSDKLSFTHTLLGQFFQGWGT 346
DB	290	AVWCHRRYFVSEYTPIDSNIAFSVNSS-----DKGEASCCDPLGAAFDCLRRMTFKGA 345
QY	347	WVASHPLTILVLSVTPVVALAAGLVPTLTDPVELWSAPNSQARSEKAFHQHGPFPFR 406
DB	346	FCVRNPTCIIFPSLAFTVCSGLVFVQVTTNPVELWSAPHQARLEKEYPDKHFGPFPFR 405

Db 862 GSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVADAGIINPFA--FYIYLTAWVSN 915
 Qy 1018 D-----RPNIKCPKGGAAAY--STSVNLTSQGVLASRFMAYHKPLKNSQD 1061
 Db 916 DPVAYAASQANIRHPPEWVHDK--ADYMPETELRIPAAEPYEAQFPFLNGLRDTSD 972
 Qy 1062 YTALRAARELANITA-DLRKVPGTDPAPFEVPPYITNVFYQOYLTLPEGLFMLSLCL 1120
 Db 973 FVEAIEKVRVICNNYTSGLSSYPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
 Qy 1121 VPTFVAVSCLLGLDLRSGLLNLSIVMLVDTVGFPAWLDWISYNAVSLINLVAAGNSVR 1180
 Db 1023 ACTFLVCAPFLNPWTAGIIVMVLATNVELFGMGLIGIKLSVPPVILIASVGIQVE 1081
 Qy 1181 FVSHITRSP--AISTKPTMLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQLIQIF 1238
 Db 1082 FTVHVALAFITAGDKN--HRAWLAEHMFAPVLQD-AVSTLLGVMLAGSEPDFIVRY 1137
 Qy 1239 FFRNLNLTLLGLLHGLVFLFVILSVG--PQVNPALALEBQ----KRAEEAAVAVMVASCP 1293
 Db 1138 FFAVLAILTVLGVNLGLVLLPVLSPFGPCPEVSPANGNLRLTPSPPEPPSVVRPAVPP 1197
 Qy 1294 NHERSVS-TADNIYVNHSPGSI-----KGAG 1319
 Db 1198 GHTNNGSDSDSYSSQTTVSGISSELRQVEAQQGAG 1234

RESULT 6
 T18538
 patched protein - chicken
 C/Species: Gallus gallus (chicken)
 C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C/Accession: T18538
 R/Marigo, V.; Scott, M.P.; Johnson, R.L.; Goodrich, L.V.; Tabin, C.
 Development 122, 1225-1233, 1996
 A/Title: Conservation in hedgehog signaling: induction of a chicken patched hom
 A/Reference number: Z18958; MUID:96205046; PMID:8620849
 A/Accession: T18538
 A/Status: preliminary; translated from GB/EMBL/DBJ
 A/Molecule type: mRNA
 A/Residues: 1-1442 <MAR>
 A/Cross-references: UNIPROT:Q90693; UNIPARC:UPI00001328B4; EMBL:U40074; NID:g13
 C/Genetics:
 A/Gene: PTC
 C/Superfamily: Drosophila membrane protein patched
 C/Keywords: transmembrane protein

Query Match 9.1%; Score 628.5; DB 2; Length 1442;
 Best Local Similarity 23.6%; Pred. No. 1.6e-35;
 Matches 284; Conservative 165; Mismatches 421; Indels 335; Gaps 47;

Qy 365 ALAAGLVFTETLTPVELSAPNSQARSKAFHDQHFQPPFRTN-QVILTAPNRSRYVD 423
 Db 111 AFAPVLAANLETNVELVWVGVRVSRRELTNRQKIGBEAMFNPLMTQTPOEDG---- 166
 Qy 424 SLLGLPKNFGSGLDLDLLLELLE--LQERLRLHQVMSPEAQRNISLDICVAP----- 474
 Db 167 -----TNVLTTLRQHLDLSAQASRVHYMYN-----RQWKLEHLCKYKSGELITEA 213
 Qy 475 --LNPNTSLYDCINSLLQYFQNNRTLLLLLTANOTLMGQ----- 512
 Db 214 GYMDQIIEVLPCLLIITPLDCFWEGAKLQSGTA--YLLGKPPQLQWINFOPLFLEELKKI 271
 Qy 513 TSQVMDKDHPLY-----CAN-----APLTF--KQGT-----ALALS--CHADY 546
 Db 272 NYQVGSWEEMLNKARVGHGYMDRPLCPNADPDCCFITAPKNKSTKPLDVALVLSSGGC---Y 328
 Qy 547 GAPVFPF-----LAIGG-----YKGKDYSAEALIMTFSL-----NNYP----- 580
 Db 329 GLSRKYMHQWQELIIGTVKNSGKLVS-AQALQTMFQMTPKQMYEHPKGYEVSHINW 387
 Qy 581 AGDPRLAQAKLWBEAFLEEMRAFQRMWAGHFQVTFTAERSLEDEINRTTAEDLPFATSY 640

388	Db	NEDKAAALEAWORMYEVVVHQSVQNSQKVLSFTT--TTDDILKSPSDSVIRVASCY 446
641	Qy	IVIFLYISIALGSSYSSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFYSYLGRSSLVILQ 700
447	Db	LLMLAVACLTM--LRWD--CAKSQGNAGLAGVLLVALSVAAAGLGLCSLIGISFNAATQ 501
701	Qy	VVPFLVLSGADNIFIVLEBYQRLPRRPGPREVHIGRALGRVAPSMKLCISLSEACPFL 760
502	Db	VLPFLALGVGDVDFLVAHAFSETGQNKRIPPFDRTGCECLKRTGASVALTISINVTAFPM 561
761	Qy	GALTPPAVTFPALTSGLAVILDFLOMSAFVALLSIDSKRQASRLDVCCC-----812
562	Db	AALIPIPALRAFSLOAAVVVVFVAVVLLIPAILSSMDLYRRDRDLFCFCTSPCVTR 621
813	Qy	---VKPO-----ELPPPGQGEGL-----827
622	Db	VIQEPQAYAEONDNICYSSPPYSSHSFAHETQITMQSTVOLRTEYDPTQAYTTTAEPR 681
828	Qy	-----LLGFFQKAYAP 838
682	Db	SEISVQVPTVTQDSLSCQSPESASSTRDLLSQFSDSSVHCLPEPCTKWTLTSTPAEKHYAP 741
839	Qy	FLHLWITRGVWLLFLALFGVLSYSMCHISVGLDQELALPKDSVLDYFLFLNRYFEVGA 898
742	Db	FLKPKAKVWVIFLGLGLLSLYGTRVRDGLDLDIVPRDTRFYDFAAQKYSF--799
899	Qy	PVYFVTTLGYNFSSEAGMNAICSSAGCNFSTFKIQYAT-----EPPEQSYLAI 948
800	Db	-----YNM-----YIVTQKADYPNVQVHLLYELHRSFSNVTYVLL 833
949	Qy	PA-----SSWYDDFDWL-----TPSS-----CCRLYISGNKD 977
834	Db	EGDRQLPKWMLHYFRDMLQGLQDAFSDWETGKITYSNYKNGSDDAVLAKVLLVQTGNRA 893
978	Qy	KFCPSVTNSLNCNKSIMTWSVRPSVEQHKYL--PWFLND-----RPN1 1022
894	Db	K--PIDISQLT--KORLVADAGIINPNA--FYIYLTAWNSNDPVYAASQANTRPHRPEW 947
1023	Qy	KCPKGGGLAAY--STSVMNTSDGOVLASGRFMAYHKPLKNQSDYETALRAARELAANITA-D 1079
948	Db	VHDK---ADYMPETRLRIPAAEPTEYQAQFPYLYNGLRETSDFVEAIEKVRACNNYTSLG 1004
1080	Qy	LKRVPGTDPAPEPYTTITNVFYEQYITLPEGLFMLSLCLVPTFAVSCLLLGLDLRSGL 1139
1005	Db	IASYENG-----Ypf-----LFWEQYIGLRHWLLLSISVVLACTFLVCALFLAMPWTAGI 1054
1140	Qy	LNLSISVMILVDTVGFPMALWDISYNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTW 1197
1055	Db	I--VVVALMTVELFGMMGLIGIKLSAVPEWVILIASVGIGVEFTVHIALAFITALGDKN--1111
1198	Qy	LERAKEATISMGSAVFACVAMTNLPGLIVLGLAKAQILQIFPFRRLNLLITLGLLHGLVF 1257
1112	Db	--RRAVLALEHWPAPVLDG--AVSTLLGVLMLAGSDFIVRVFFAVLAILITLGVNLGLV 1169
1258	Qy	LPVILSYVG--PDVNPALAA---LEQKRAEEAAVAVMWASCPNHPSPRSTADNIYVNHHSFE 1312
1170	Db	LPVLLSPFGPYEVSACGRNRLPTPSPPEPPIVRFPALPGHTNNGS--DSSDSKYSQ 1227
1313	Qy	GSIGK 1317
1228	Db	TTVSG 1232

RESULT 7
S44797
F09G8.4 protein - *Caenorhabditis elegans*
C.Species: *Caenorhabditis elegans*
C.Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C.Accession: S44797
R.Anderson, K.
submitted to the EMBL Data Library, February 1993
A.Description: Sequence of the C. elegans cosmid F09G8.
A.Reference number: S44796

A:Accession: S44797
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-915 <AND>
A:Cross-references: UNIPROT:P34389; UNIPARC:UPI000013B859; EMBL:L11247; NID:g156280; PTL:11247
C:Genetics:
A:Introns: 51/1; 214/3; 255/3; 326/3; 382/2; 409/3; 455/3; 505/3; 569/2; 702/3; 836/3

Query Match 8.7%; Score 600; DB 2; Length 915;
Best Local Similarity 22.7%; Pred. No. 8.9e-34;
Matches 236; Conservative 158; Mismatches 422; Indels 224; Gaps 37;

QY 10 LLWALLRLAQSBPYTTIHQPGYCAFYDCKGNPELGSMLTSLNSVCSLSNTPARKITGD 69
DB 14 VLFLLLHLALCO-----AKVM-TECDGEDSNHPPCKTKNTYLPIITVTRSLNPT 64
QY 70 HLILLQKICRLTYTGN-TQACSAQLVSLASLITKALLTRCPACSDNFVNLCHNT 128
DB 65 YMAREFYKSYLVQEBEDKAQVCCTELQKGMTDRISNAATILGSCSPDFNFAKWCQFT 124
QY 129 CSPNQSILFINTRVAQLGAQLPAVVAIFYQHSFAEQSYDSCSRVRVPAATLAVGTM 188
DB 125 CSPDQSKFMKMET-----TCPKNVVVMEFKVNRDFVEGLYESCRHTWFANGALRLMSL 180
QY 189 CGVYGSALCNAORWLNFQDGTG-NGLA---PLDITPHLEPGQAVGSGTOPLNEGVARCN 244
DB 181 GGRKVS-----FENFYFGMTKNLAQSIPIINTEPQFSRMQNMNIPTTP-----CH 225
QY 245 ESQDDVATCSQDC---AASCPAIARPOALDSTFVILGQMPGSLVLIIILCSVPVVTL 301
DB 226 KSAQPKVPACGALDCPTNAHQLVDSKEVHLGKTVFHPHP-DFEWLLKICGLA-LTVL 283
QY 302 LVGF-----RVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFPQG---WGTWVAS 350
DB 284 LVFLTKYSCHRRSAPNGEDCYVDLKGK-----LEVQPEGLCARYANAVIK 330
QY 351 WPLTILVLSVPPVVALAG-LVTFELTITDPEVLSAPNSQARSEKAP-HDQHGPPPTRN 408
DB 331 HPLIFVSLGLIVAAACCSGNFKHSLTHSDVQSAADGETTRNEKKEPIHS--PGPNHRIE 388
QY 409 QVILITAPNRSSRYDSSLGPKNPSGILDLDLLELELQERLRLHQLVQVSPQAQRNLSLQ 468
DB 389 QIFNLPTT-----KSMFNPLPEEMFQLVGNIGNLT--ACYGNSVVKLD 432
QY 469 DICVAPLNPDNTSLYDCINSLQYFQNNRTLLLLTANQTLMGQTSQVDWKDHPLYCAN 528
DB 433 DICYKPIGKN---HGCAIMSPTNYFQ----- 455
QY 529 PLTFKQGTALALSCHWADYGAPVFFFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQ 588
DB 456 -----YYTARTIMITVLRG-----PE-DQ 474
QY 589 AKLMEEAFLERMAFQRRMAGMFQVTFEARSLEDEINRTAED--LPIFATSVIVFLPY 646
DB 475 AIANETAFLNMMSRYEMKHANF--TFWTETSVAAEIIHTAVETDKIVSVIACAAVLWVI 531
QY 647 ISLAGSY--SSWRGVWDVKATLGLGVAVVLGAVMAAGFFSVLGRSSRLVLQVVP 704
DB 532 TMLGINHPPESSILSALVHHKLLISISAMVISVIVSMCSIGMFSLFGVHATDNAVILFF 591
QY 705 LVLSVGADNIFIFVLEYQRLPRRPEP-----REVH--IGRALGRVAPSMCLCSLEATCF 758
DB 592 VITCLGINRIFVIRTFQANGCHYGLPNI SYREMHNRISNVMRRSIPVLTNSLICSTCL 651
QY 759 FL-GALTP-----MPAVRTFALTGSLAVILOFLQMSFAVALLSDSKQ----- 802
DB 652 FLAGGVLPVSVSNPAVEVFARHAGLALMDTAFYLLVMLFLFQYDARREMSGKCEINPW 711
QY 803 ----EASRLDVCCKVPQELPPQGGEGLLGFFOKAVAPFLHWHITRGVLLFLALFG 858
DB 712 YELSNESKINLCWEAVDGNLRSP-----VDWFKLATAPLLKKICHIWATFFVSLI 764
QY 859 VSLYSGMCHISVGLDQELALP-----KDSYLLDYFLFLNRYFEVGAPVYFT 904

```

Db      765 IACYCTLCLEFGNQMAFSEVCIFLPSHLNGSFQSYLTUKFQNNENLNGPPLWFFV 824
Qy      905 TLGYNFSSBAGNAICSSAGCNFFSTQKIQ---YATEFPEQSILAIPASSWVDDPIDWL 961
Db      825 EGDVWKHDPKMQNKFCCTLAGCDNSMGNKIRSLAYAEVY-KGNYLHGDVNIWLDVLOPM 883
Qy      962 TP-SSCCRLYISGPNKDKFC 980
Db      884 HPRGSCCKM-----DGKQFC 898

RESULT 8
T25600
hypothetical protein C32E8.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25600
R:Gattung, S.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C32E8.
A:Reference number: Z20056
A:Accession: T25600
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-933 <GAT>
A:Cross-references: UNIPROT:P91129; UNIPARC:UPI00000611CF; EMBL:U88308; PIDN:AA842325.1;
A:Experimental source: strain Bristol N2; clone C32E8
C:Genetics:
A:Gene: CESP:C32E8.8
A:Map position: 1
A:Introns: 145/1; 177/2; 385/1; 418/3; 833/3; 878/3

Query Match      8.7%; Score 598.5; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 1.2e-33;
Matches 251; Conservative 181; Mismatches 362; Indels 303; Gaps 37;

Qy      341 PGGWGTWASWPLTILVLSVIPVVALAAGLVFTELTTDPVLMWASPNQARSEK-AFH-- 397
Db      14  FROLGLICDHLPFPVFPPLPFTAAMGVGLHNLPLSDAVLYLFTPLGAQSKMERMSIHEK 73

Qy      398 ----DQFGP---PFRNQVILTPNRSRYRDSLLGKPNFSGILDLDLLELELQER 450
Db      74  WPLTDNNYIPGRAVTSQREIQVTALARN----DSNILDPKFANAVYQLDKY-----IQTR 124

Qy      451 LRHLQWSPQAQRNISLQDICVAPLNPDTSLDYDCINSLLQYFQNNRTLLLTANQITLM 510
Db      125  VAVLH----- 129

Qy      511 GQTSQVDWKDHFYLCANAPLTPKDGTA-----LSCMADYG-----APVPPFLAIGGY- 559
Db      130 -----NGHYYSYKNCLOQYKNGCPSGNKHVHLSLDLHNGFNITPYPRFGSEGGYI 181

Qy      560 -----KGDYSE-----REALIMFSLNYP-----AGDPLAQAKLWEAFLEE 599
Db      182 GSSLGGVTVKGENETDILASAKAFWMIYHLKFPHEMSYISGE-----WE---LEL 230

Qy      600 MRAFORMAGMP-QVTETAERSLEDEINRTTAEDLPFATSIVIVFLYISLALGSY--SS 656
Db      231 GMLTQYEDPDVISITYFHSITQTLABELKKNADTLIPRFIISITLIVFSLTCLSLFIDGS 290

Qy      657 MSRVNVDKATGLGGVAVVLGAVNMAAMGFFSYLGIRSLVILQVVPFLVLSVGDADNIFI 716
Db      291 PSIDVWLSKPIILSVLSGAGTALTGVGFLSMGMPYN-DIVGWMPFLVLAAGVDNMF 349

Qy      717 FVLEVQRLPRRGEPREVH--IGRALGRVAPSNLCSLSEACFPFLGALTTPMAVTRTAL 774
Db      350 MVAAV-----RRTSRTHVHERMGECLADAASLITSTSDVLSFGVGALTTPAVQIFCV 405

Qy      775 TSLGAVILDFLLQMSAFVALLSLDSKQEAASR-----LDVCCVK----- 814
Db      406 YTGVAIFPAFYQITFFAACLALAMKHEASGRNSLFLTEAVSAEKKTSLSSTPQRLFNLGS 465
```

```

Qy      815 -PQELPPPGQGGGLLGGFPQKAYAPFLHLHWITRGVVLLLP-LALFQVSLYSMCHISVGLD 872
Db      466 VPDHSHSDVKQPLTSRPFGEWYAPVLMHPVVRGIAWVFWIYLLGAS-YGCSRIKEGLE 524

Qy      873 QELALPKDSYLDYFLFNRYF-EVGAPVVFV-----TTLGYNFSS 913
Db      525 FVNLIVBDSYALPHYRLLEKIFWKYGOQVQVINNAPDLRNHTSRDRVHMLVDFATSKH 584

Qy      914 A-GMNAICSSAGCNFFSTQKIQYATEFPEQ-----SYLAIPASSWVDDPIDWL 961
Db      585 AIGMESV-----QFWLPEMERYQKELEVOIIDSSFYGLLHHLFLASKTNPLAEDIYW- 637

Qy      962 TPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVQEFHKYLPWFLNDRN 1021
Db      638 -----GMPD----- 642

Qy      1022 IKCPKGLAAVSTSVNLTSDGOVLAS-RFMAVHKPLKNSQDYTEALRAARELANITADL 1080
Db      643 -----DDNGTWVKSFRFILGMDLVTTMOQTDATMSPREVAARW----- 681

Qy      1081 RKPCTDPAFEVPPYITINVFYEQYLITLPEGLFMLSLCLVPTFAVSCILLGLDLRSGLL 1140
Db      682 -----PEFNVTTPMPITWMTDQYIIIPNTVQNIILVMIIVIAVLFIQPMCS-LW 733

Qy      1141 NLLSYMILVDTGVFMALWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db      734 VALACASIDFGVIGYITLWGNLDAISMITIIMSIGFSDVYSAHIAYGTVVSRDPTAAGR 793

Qy      1201 AKEATISMGSAVFAGVAMTNLPGIILVGLAKAQLQIPEFRNLNLTLLGLAHGLVFLPV 1260
Db      794 VKEALSALGWPLSQG-AMSTIIVASVLADIPAYMI-VTFEKTVVLVLSLGLHLGLVFLPV 851

Qy      1261 ILSYV-----GPDVNPALALQKRAEEA-----VAAMVVASCPNHPRSVS 1300
Db      852 LLSIFVPGCCIIIPSPHPGHPSAQKIEQIRINAISSSLDLDTVAPLRASSPISPHRLE 911

Qy      1301 -TADNIYNHSPGSIK 1316
Db      912 YDESPTVHNRSKNSIK 928

RESULT 9
T13952
membrane protein ptch2 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T13952
R:Motoyama, J.; Takabatake, T.; Takeshima, K.; Hui, C.
Nature Genet. 18, 104-106, 1998
A:Title: Ptch2, a second mouse Patched gene is co-expressed with Sonic hedgehog.
A:Reference number: Z17830; MUID:98122566; PMID:9462734
A:Accession: T13952
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-1182 <MOT>
A:Cross-references: UNIPROT:Q35595; UNIPARC:UPI0000001587; EMBL:AB010833; PIDN:BA24691
A:Experimental source: strain BALB/cCrSlc
C:Genetics:
A:Gene: ptch2
C:Superfamily: Drosophila membrane protein patched
C:Keywords: transmembrane protein

Query Match      8.4%; Score 583.5; DB 2; Length 1182;
Best Local Similarity 24.2%; Pred. No. 1.8e-32;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

Qy      330 SFSTHTLLGQ-----PFGQ-----WGTWVASWPLTILVLSVIPVVALAAGLVFTEL 376
Db      20  SSAPHILAGSLQAPLWLRAYFQGLFLSLGCRIQKHGKVLFLGLVAFGLALGLRVAVIE 79

Qy      377 TDPVLMWASPNQARSEKAFHDQHPG-PFRFNQVILTPNRSYRYSLLGLGPKNFSGI 435
Db      80  TDLEQLWVEVGSVQELHYTKKLGEBAAVTSQMLI-----QTAHQEGGNVLTPE-----A 131
```


QY 996 ITWGSVRPSVEQPHKYLPHFLNDRPNKCPKGLAAYTSVNLTSQGVLT-ASRFMAYHK 1054
DB 630 -----VFLSQPOF-----AKYNRDVLVLTEDGEHLEASRI----- 658
QY 1055 PLK-----NSODVTEALRAARELAANITADLRKVPVGTDPAPVFPYITITNVEYEOYLITLP 1110
DB 659 PVLRLHVGANGSRANLFRRLAE--TSELQ-----TGVYADPFO-----FAEQYNVAVLP 706
QY 1111 EGLFMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTVGFMALWDISYNVSLIN 1170
DB 707 GTLSSIAYAGVAVVSLILPEVAS-LWVSFSIVSINIGILGFMTFWSVRLDPTSMVT 765
QY 1171 LYSVAGMSVEFVSHITRSFAISTKPTWLSRAKEATISMGSAVPAGVAMTNLPGLVGLIA 1230
DB 766 IVMSIGFCVDFAHLAYNAFKAGQNMDSRMRNALYAVGAPILMS-ATSTIIIGVSPMASA 824
QY 1231 KAQLIQIPFRLNLLITLGLLHGLVFLPVLS--VVGPDVNPALALEOKRAEEVAAV- 1287
DB 825 ESYVFR-SFKTIVMLVILGALHGLVLPVLLSMFYCGG-----SSKAKEHIDAVD 875
QY 1288 -MVASCNPHPSRVST 1301
DB 876 QKLQAYNNPARTAS 890
RESULT 11
Ti8291
patched protein - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18291
R:Concordet, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.;
Development 122, 2835-2846, 1996
A:Title: Spatial regulation of a zebrafish patched homologue reflects the roles of sonic
A:Reference number: Z18860; MUID:96379744; PMID:8787757
A:Accession: T18291
A>Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1220 <CON>
C:Cross-references: UNIPROT:Q98864; UNIPARC:UPI00001328B3; EMBL:X98883; PIDN:CRA67386.1
C:Genetics:
A:Gene: Ptcl
C:Superfamily: Drosophila membrane protein patched
Query Match 8.3%; Score 572; DB 2; Length 1220;
Best Local Similarity 22.4%; Pred. No. 1.2e-31;
Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;
QY 321 KGTSLSDKLSFSTHTLLGQFPQGWGTWASWELTILVLSVIPVVALAAGLVFTELTPDV 380
DB 51 KGKAVGQKAPLWTRARFOAFLPSLGCHIQRHCGKVLFIGLLVFGALSVGLRVAATHTDE 110
QY 381 ELWSAPNSQARSEKAFHDHFG-PFPRTHQVILTAPNRSSRYSDLSLLGPKNPSGILLDUD 439
DB 111 KLWVEAGSRVSKELRYTKEQGEESVFTSQMLIQTPKQEG-----TNILTOE 157
QY 440 LLELLELERLRLHQLQVSPQARNISLQDICVAPLNP--DNT-----SLYDCINSL 490
DB 158 AL--LLHLEAALSASKVQSVLSYKGSWDLNKKICFKGVPIENVMTERMDKLPFCMVTP 215
QY 491 LOYFQNNRTL-----LLLTANO--TIMGQTSQVDKDHFLYCA 526
DB 216 LQCFWEGSKLOGSAVLPQMPDIPQWNLDPKLMBELSQTSLEGFREMLDKAQGVHAYM 275
QY 527 NAPLTFKDTALALSCHADYDAP-----VPPFLAI---GGYKG----- 561
DB 276 NRPCLDPSDT-----DC--PHSAPNKDPWQVPNIAELQGGCHGFSKPFKMHQOEELIGR 329
QY 562 -KD-----YSEBALINTFSL-----NNYPAGDPRLAQAKL-----WEBAFLTE 599
DB 330 VKDSQNALQSALQTMFLMSPKQLYEHFKDDYELHNNWEDKATALLSQRKPFV 389
QY 600 MRAP--QRRMAGMFQVTTAERSLEDEINRTTAEDLPFATSYVIFLYISIALAGSYSSW 657

DB 390 VHGISIPONSSSNVYAFSTT---TLNDIMKSFSDSVIRVAGGYLLMLAYACVTM---LRW 443
QY 658 SRVWDSKATLGLGGVAVVLGAVMAAMGPFYSLGRSSLVILQVVPFLVLSGADNIFIF 717
DB 444 D--CAKSOQAVGLAGVLLVALSVAAGLGLCSLGLSFNAATQVLPESLAGIGVDDMFL- 500
QY 718 VLEYQRLPRRPGEPREVIHTRALGRVAPSMGLCSISEAICFFLGAITPMPAVRTFALTSG 777
DB 501 -LGHSTFETRSNIPKERTGDCRLRTGTSTVALTSVNMIAFFMAALVPALRAFSLQAA 559
QY 778 LAVILDFLOMSAFVALLSLDSKQBSASRLDVCCC-----VKQOE----- 817
DB 560 VVVVFNFMALLIFPAILSLDLRREDKELDILCCFYSPCSSRVTIQIQEQLSDANDNHQ 619
QY 818 -----LPPPGQ----- 823
DB 620 RAPATPTTGTSTITTTTITTTTQVATQDAAGQHIVTLPPTSQISTTPSPSVLSTPTTP 679
QY 824 -----GEGE-----LLGFFQKAYAPFLHWTIR 846
DB 680 TTDYGSQVFTTSSSTRDLLAQVERPKGREGCVLPFPFRNLLSSFAREKYAPLLKPEIK 739
QY 847 GVLLLFALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTL 906
DB 740 TVVVVVFALLSLSYGTITMVDHGLYITDIPRDTQEYEFITAQFKYFSF-YNNVLVTMD 798
QY 907 GYNFSSEAGMNAICSSAGCNCNPFQKIOYATFEPQSVLA-----IPASSWDDPDIM 960
DB 799 GFDIAR-----SQRLQLQHNAFNSVKYVVDGNHKLGP-RMWLHYFQDM 841
QY 961 LTFSSCC-----RLYISGPNKDKFCPSTVNSLNCNKC 993
DB 842 LKGLQATPDADWEAGKITVDSYRNGTDEGALAYKPLIQTSKKEPNYSQTSRRLVDG- 900
QY 994 MSITWGSVRPSVEQPHKYL-PWFLNDRPNKCPKGLAA-----YST--- 1034
DB 901 -----DGLIPPEV--FYIYLVTVWSND-----PLGYAASQANFYPHPREWIHDKYDTGE 948
QY 1035 SVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAAN-ITADLRKVPGTDPAPVFP 1093
DB 949 NLRIFAAPLEPAQPFYINGLURQASDFIEATESVRTICEEFMRQIKPNYG-----Y 1002
QY 1094 PYTITNVEYEOYLITLPBGLFMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTV 1153
DB 1003 PF-----LFWEOYIGLRHWFLLSISVVLACTFLVCAILLNPNWTAGVI-VFILPMVTVELP 1057
QY 1154 GPMALWDISYNVSLINLSVAGMSVEFVSHITRSF--AISTKPTWLSRAKEATISMGSA 1211
DB 1058 GIMGLIGIKLSAIPVVPVILLIASVGIVFTVHIALGFLTAIGDRNT---RSAVAMEHMFAP 1114
QY 1212 VPAGVAMTNLPGLVILGLAKAQLIQIPFRLNLLITLGLLHGLVFLPVLSVGP--DV 1269
DB 1115 VTDG-AISTLLGLVLMAGSEDFIMRYFFAVAILTLLGILNGLVLLPVLMSGPPAEV 1173
QY 1270 NPA 1272
DB 1174 VPA 1176

RESULT 12

T29590

hypothetical protein F55P8.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T29590

R:Gattung, S.; Scheet, P.; Kemp, K.

A:Description: The sequence of C. elegans cosmid F55P8.

A:Reference number: Z20647

A:Accession: T29590

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A;Residues: 1-889 <GAT>
A;Cross-references: UNIPROT:P91346; UNIPARC:UPI000017BA20; EMBL:U80447; PIDN:AAB37812.1;
A;Experimental source: strain Bristol N2; clone F55F8
C;Genetics:
A;Gene: CESP:F55F8.1
A;Map position: 1
A;Introns: 36/3; 66/3; 98/3; 149/2; 200/3; 240/2; 279/3; 358/1; 481/2; 569/3; 606/3; 697/3

Query Match 7.8%; Score 539.5; DB 2; Length 889;
Best Local Similarity 23.0%; Pred. No. 1.5e-29;
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;

QY 401 FGPFFRTNQVIL-----TAPNSSVRYDLSLLGPKNFSGILDLILLLELLE 446
DB 31 FPGYSYTERRIIHDAWPLVDGTFVAGRAVTS--REVQVAVVARSGGNLDVRSNELKL 88
QY 447 LOERLRHLQVMSPEAQRNLSLODICVAPLNPDNTSLYDCCNSLLQYFQNNRTLLILLTAN 506
DB 89 MESFIRN-NITVQFSNRWTSFADLCIA--GPDGR-----CANNDHIQ----- 127
QY 507 QTLMGQTSQVWDKDFLYCANAPLTFKDGDTALALSCMADYGAPVPPFLAIGYK----- 560
DB 128 --LASRLHQ-----HGINTITYTVRLSDKSAVIAS-----ALGGVKLAKGDN 167
QY 561 GKD-YSEAEALIMTSLNNYP-----AGDPLAQAKLWEAEFLAEAFQRMAGMFOV 613
DB 168 GENIIVEATAWLLIYQLKFPNEISYVSG-----LWEREFKNKMDY-KQAKYIS 218
QY 614 TTAERSLEDEENRTAEDLPFPATSYIVFLY-----ISLAGSYGSWSRVMVDSKAT 667
DB 219 TYFHSQTSLDELRNARLAPKFIQAFVILCVFSVLCSIVTIKSGYIDW----VVTXPI 274
QY 668 LGLGGVAVVLGAVMAAGFFSYLGIRSSLIIVQVFPFLVSLVGADNIFIVLEYQRLPRR 727
DB 275 LSVLGVSNAGMGASANGMLTYLEIQYN-DIIAVMPLVAVGTDMFMELVASLKRTR- 332
QY 728 PGEPREVHIGRALGRVAPNMLCSLSEALCFPLGALTMPAVRTALTSLGLAVILDLLQ 787
DB 333 -NLKYDQRIAEACMADAASILITALTALDLSFGVTGTTTIPAVQIFCIYTMCAILLTFAVQ 391
QY 788 MSAFVALLSDSKQRE-----ASRLDV--CCCVKQPQLPPPGOG----- 824
DB 392 LTFEFCALLVYTRIESQGLHSLWLRPAVYTSSTPLNVKLFWLGSQPK-PLPSCGTVSS 450
QY 825 -----EGLLGFFQKAYAPFLH-WITRGVLLFLFLAFGVSLYS 863
DB 451 TSSVSTWTSQATSPASKHLHCAATSFFRNWYAPVLMQPMI-RAIAGLWLYLIGISIYG 509
QY 864 MCHISVGLDQELALPKDSYLLDYFLNRY-YPFVGAPVYFVTT-----LGVNFSSE 913
DB 510 CTHLKEGLEPANLLVDDSYATPHYRLKHYWHYGASLQIVVSNPPDLRDPVERINMDKM 569
QY 914 AGMAICSSA-GCNSFSF-----TOKIQYATE-----FPEQSYLAIPASSWYDDFI 958
DB 570 ASTFANCKVAIGDSQVFWLREMQVSEEHKIQYDNEKFFYDHAAQYISDMSQPWVVDV 629
QY 959 DWLTPSSCCRLYISGPNKDKFCPSTVNSLNLCLNCKMSITWGSVRPSVEQPHKLPFLND 1018
DB 630 -----WGRNN 634
QY 1019 RPNIKCPKGLAAYSTVNLTSQGVLASRPMAYHKPKNSQDYTRALRAARELANITA 1078
DB 635 -----NSERIKTFRPMIGMRDISITTTQTEATNTFRFIAISRP-- 672
QY 1079 DLKRVPGTDPAFEPVFTITNVFQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
DB 673 -----BOYNVTTTWPLWFLFDQVALVVPNTMQDIIIVAVCMVLISALLIPQPCVSF 723
QY 1139 LLNLLSIVMLVDVTGFMALWDISYNAVSLINLVSAGMGSVEFVSHITRSFALSTPTWL 1198
DB 724 WV-AVTIGSIDLGVLFMTLNNVNLDAISMITIIMSVGFSVDYSAHIYAYVISKESTTS 782
QY 1199 ERAKEATISGSAVFAGVAMTNLPGLVILGLAKQLIQIFFFRNLNLLITLLGLHLGVFL 1258

DB 783 ARVCDALDGLGWPAQG-AMSTILAVSLSDVPAYMI-VTFEKTVFLAISIGFLHGLVFL 840
QY 1259 PVILS-YVG 1266
DB 841 PLMLSVFVG 849

RESULT 13
T27969
hypothetical protein ZK675.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T27969
S;Sims, M.
submitted to the EMBL Data Library, November 1994
A;Reference number: Z20448
A;Accession: T27969
A;Status: preliminary; translated from GE/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1405 <WIL>
A;Cross-references: UNIPROT:O09614; UNIPARC:UPI00000835A5; EMBL:Z46812; PIDN:CAA86843.1;
A;Experimental source: clone ZK675
C;Genetics:
A;Gene: CESP:ZK675.1
A;Map position: 2
A;Introns: 31/3; 59/3; 100/3; 244/2; 485/1; 950/2; 1298/2; 1335/3
C;Superfamily: Drosophila membrane protein patched

Query Match 7.4%; Score 514; DB 2; Length 1405;
Best Local Similarity 20.2%; Pred. No. 1.7e-27;
Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSFSTHTLGCQFPQCHGTWY--ASWPLTILVLSVIPVWALAAGLVPTELTD 378
DB 100 RKGATGNRYALYSRLIQKLFALGNTVHRNAWSI-ILAVSMIFAVC-CYGLQYVHIETD 157
QY 379 PVELWSA-----PNSQARSEKAFHDHFGPPF-RTN-----QVILTAPNRSS 419
DB 158 IVKLWVAOGRLDEELNPLNFKAMRNVTGDS--GPELPRENGLGGGVQVLIQTPEYEG 215
QY 420 YRYDSLLGPKNFSGILDLDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 472
DB 216 --QDALAAGP-----LLKHVEIMKHIASEFNVSVHGVDMWSLSDICF 253
QY 473 APLNPD-----NTSLYD-----CC-INSILQYFQNNRTL-----LIL 503
DB 254 KPAPSPVAADSAASSLGVDIKIVPCWIPTIDCFWEGSKALGPHPSLPKSSLGLPLGMLL 313
QY 504 TANQTLMGQTSQVWDK-----HFLYCANAPLTF----- 532
DB 314 SS-----LSDGDMIRWSDPDPIAVIDEIHRSFNLGSHYTPFERAGVSHGYMDRPFCDPLDP 369
QY 533 -----KDGATALASCADYCAPVFPFLAIGYKGYKDYSEAE 568
DB 370 ECPPMKAKYFDVCPHIDRVREIAKYGTELEBEKKKSGYSFFDEL---GRKKREAGDQ 426
QY 569 ALI----- 571
DB 427 KMIHPAQPADSIPTIEDVAPQVPVSTAPIPTTTTSLPEEARAAEKEKKQKARELKDYC 486
QY 572 -----MTFSLNNYP-----AGDP 584
DB 487 KSVKRSAPFELWKKKQKWEPEVMSNNYPQNVDAAEWTCGCGSGFASNVNLNWPEDMTLGNP 546
QY 585 RLQA--AKL-----NEEAFLEB-MRA 602
DB 547 RRKKGKGLSGDALQSFLVASPADVFLRFKQKPGNRNKNKTGLDMDAWNETHAAEQVLOA 606
QY 603 FQREWAGMFOVTTAERSL-----EDEINRTTAEDLPIFATS-----YIV 642
DB 607 WQR-----NFT--KSLYNHKNVDEEDGNERETLH--PLASTSIADMLEBFCQFNNTI 654

[illegible]


```

Db      857 MVTIDIGVIGLSVSKVLDPISMITIMSGSIFBSAHITHGFVNSDLSAFDRCDV 916
Qy      1204 ATISMGSAVAVGAMTNLPGLVILGLAKAQIQIFFFLNLLITLLGLLHGLVFLPVLS 1263
Db      917 AMEKLANPVG-SLSTILGVFLAFIDSVYVLPFFKTSISLV-LIIGAWHALMELPILLS 974
Qy      1264 YVGP-----DVNPALALEQKRAEAAVAVVASC 1293
Db      975 MCPIVIERLSDASKKASDRRRKLNKNSVAINLP 1010

RESULT 15
S06119
membrane protein patched - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: S06119; A33468
R:Nakano, Y.; Guerrero, I.; Hidalgo, A.; Taylor, A.; Whittle, J.R.S.; Ingham, P.W.
Nature 341, 508-513, 1989
A:Title: A protein with several possible membrane-spanning domains encoded by the Drosophila
A:Reference number: S06119; MUID:90015164; PMID:2797178
A:Accession: S06119
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-1299 <NAK>
A:Cross-references: UNIPROT:P18502; UNIPARC:UPI000014EA18; GB:X17558; NID:98389; PIDN:CA
R:Hooper, J.E.; Scott, M.P.
Cell 59, 751-765, 1989
A:Title: The Drosophila patched gene encodes a putative membrane protein required for se
A:Reference number: A33468; MUID:90058658; PMID:2582494
A:Accession: A33468
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-110,'R',112-273,'G',275-331,'R',333-635,'P',637-861,'DVF',878,'Y',880-1299
A:Cross-references: UNIPARC:UPI000016BCA3; GB:M28418; GB:M28999; NID:g552097; PID:g55209
C:Genetics:
A:Gene: FlyBase:ptc
A:Cross-references: FlyBase:FBgn0003892
A:Map position: 2 44D3-D4
C:Superfamily: Drosophila membrane protein patched
C:Keywords: glycoprotein; transmembrane protein
F;74-92/Domain: transmembrane #status predicted <TM01>
F;427-448/Domain: transmembrane #status predicted <TM02>
F;456-503/Domain: transmembrane #status predicted <TM03>
F;529-555/Domain: transmembrane #status predicted <TM04>
F;557-585/Domain: transmembrane #status predicted <TM05>
F;677-699/Domain: transmembrane #status predicted <TM06>
F;967-1017/Domain: transmembrane #status predicted <TM07>
F;1019-1047/Domain: transmembrane #status predicted <TM08>
F;1061-1086/Domain: transmembrane #status predicted <TM09>
F;1093-1121/Domain: transmembrane #status predicted <TM10>
F;142,298,335,388,807,861,1194,1271/Binding site: carbohydrate (Asn) (covalent) #status

Query Match      6.78; Score 464.5; DB 1; Length 1299;
Best Local Similarity 20.48; Pred. No. 4.4e-24;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

Qy      321 KGTSLSKLSPSTHTLQGFQCGMTWVASWPLTILVSLVPVVAALGVFTLFTDPV 380
Db      43  KGKARGSRITAIYLRVSFQSHLETGLGSSVQKHAGKVLFAVILVSTFCVGLKSAQIHSKVH 102
Qy      381 ELWSAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSRYVDSLLGLPKNFSGIL 436
Db      103 QLWTIQEGGLEAEIAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy      437 DLDLLLELLEQLERLHLQVNSPEAQNISLQDICYAPLNPNTSLY-----DCC 486
Db      154 HLEVLVATAVKVLHYDTE-W-----GLRDNCNPFSPFEGIIYIQLRHLPDS 204
Qy      487 INSLQVFN-----NRTLLLTAN-----QTLMGQTSQVDKDH 521
Db      205 IITPLDCFEWGSQLLGPESAVVIGLQRLNLTWTLNPNASVMQYMKQMSSEKISFDFETV 264

```

Search completed: April 7, 2006, 19:22:06
 Job time : 63 secs

```

Qy      522 FLYCANAPLTFKDGTTALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db      265 EQMKRAATA---SGMKEPCLNPLNPNCPDTAPNKNSTOPPDVGA-----ILSGCGVGYA 317
Qy      561 GKDYSEAEALIMTFSLNNYPAGDPLRAQA-----589
Db      318 AKHMHPPEELIVGAKGN-RSGHLRKAQALQS VVQLMTEKEMYDQMDNYKVHHLGWTOE 376
Qy      590 -----KLWEEAF---LEENRAFORRMAGMFOVTTFAERSLEDEINRTTAEDLPPIFATSY 640
Db      377 KAAEVLNAWQRFNSREVEQLLRKQSRITATNYDIYVSSAALDDILAKFSHPALSIVIGV 436
Qy      641 IVIFLYISLALGSYSSRMVWDKATLGLGGVAVVLGAVMAAMGFFSYLGISSVLQ 700
Db      437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLCMCFSTAAGLGLSALLGIVENAASTQ 492
Qy      701 VVPFLVLSVAGADNIFIFVLEYQRLPRPGBPREVHIIGRALGRVAPSMLLCSLSAICFFL 760
Db      493 VVPFLALGLGVDFHIFMLTAAYAESNRR--EQTKL-----ILKKVGPSTILFSACSTAGSFFA 546
Qy      761 GALTTPMAVRTFALTSLGLAVILDLLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
Db      547 AAFIVPALKVFCLOAAIVMCSNLAAALLVPPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
Qy      815 POELPP-----PGQEGEGL-----LIGFF 832
Db      607 PKVAPVPLVPLNNNGRGARHPKSCNNRVALPAQNPLEQRADIPGSSHSLSASFSLATFA 666
Qy      833 QKAYAPFLHWHITGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDFELFLNR 892
Db      667 FQHTVFLMRSWKFLVTVMGFALAALISLYASTRLQDGLDIIDLVPKDSNEHKLDAQTR 726
Qy      893 YEEVGAPVYVTTTLYGYNFSSAGNAICSSAGCANNFSFTOKIQVATEPPEQSYLAIP--- 949
Db      727 LF--GFTSYMAVTOG-----NFEYPTQQLLDDY-HDSFVRVPHVI 764
Qy      950 -----ASSWVDVDFIDMLTSSCCRLYISGPNKDKFCFSTVNSLNCUNK----- 993
Db      765 KNDNGGLPDFWLLLFSEWL-----GNLQKIPDEEYRDGRLTKECFWFPNASSDA 812
Qy      994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db      813 ILAYKLIYQGTGHVDPNDPKELVLTNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTL 868
Qy      1029 LAA--YSTSVNLTSDGOVLASRFMAYHKPLKNSQD-----YTEALRAAR 1070
Db      869 LRANCIRNRANGASQCKLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQMPPFYLHGLTDT 926
Qy      1071 ELAANITADLRKVPCTDPAFEV--PPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSC 1128
Db      927 QI-KTLIGHIRDLSVKYEGFGLPNYPSPGIPFIFWEQYMT-LRSSLAMTLACVLLAALV 984
Qy      1129 LLLGLDLSGLNLLISVIMLVDTYGVFWALWDISYNAVSLINLVSAVGMSVFEVSHITS 1188
Db      985 SULLLSWAAVILSVLSASTLAQIFGAMTLLGKLSAIPAVILSVLGVMLCF--NVLLS 1042
Qy      1189 FAISTKPTWLERAKATISMSGSAVPAVAMTNLPGILVLGLAKAQIQLIFFFLNLLITL 1248
Db      1043 LGFMTSVGNRRQVQLSQMSGLPLVHGLMTSGVAVFMLSTSPFEFVIRHFWCLLIVLVC 1102
Qy      1249 LGLLHGLVFLPVLSYVGPDVNPALALEQKRAEAAVAAVWVASCNPHSPSVSTADNIYN 1308
Db      1103 VGACNSLLVFPILLSNVGPEAE-LVPLE-----HPDRISTPSPPLPVR 1143
Qy      1309 HS 1310
Db      1144 SS 1145

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:13:14 ; Search time 265 Seconds
(without alignments)
3546.278 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGATSNFLPNNRGQF 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6872.5	99.5	1359	1 NPCL1 HUMAN	Q9uhc9 homo sapien
2	5421.5	78.5	1331	1 NPCL1 RAT	Q6t3u3 rattus norv
3	5407	78.3	1333	1 NPCL1 MOUSE	Q6t3u4 mus musculus
4	3179	46.0	1132	2 Q4T749 TETNG	Q4t749 tetraodon n
5	2413	34.9	1277	2 Q9JUG3 CRIGR	Q9j1g3 cricetus
6	2412	34.9	1277	1 NPCL1 PIG	P56941 sus scrofa
7	2405	34.8	1276	2 Q8MKD8 FELCA	Q8mkd8 felis silve
8	2404.5	34.8	1276	2 Q9GK52 CANPA	Q9gk52 canis fami
9	2402.5	34.8	1278	1 NPCL1 HUMAN	O15118 homo sapien
10	2402	34.8	1289	2 Q59GR1 HUMAN	Q59gr1 homo sapien
11	2400	34.7	1277	2 Q7TMD4 MOUSE	Q7tmd4 mus musculus
12	2399	34.7	1276	2 Q9NQO0 FELCA	Q9nqo0 felis silve
13	2395	34.7	1276	2 Q8MI49 FELCA	Q8mi49 felis silve
14	2385	34.5	1278	1 NPCL1 MOUSE	O35604 mus musculus
15	2367.5	34.3	1277	2 Q9GJC9 BOVIN	Q9gjc9 bos taurus
16	2361	34.2	1286	2 Q9T775 RABIT	Q9tt75 oryctolagus
17	2338	33.8	1209	2 Q4RWY5 TETNG	Q4rwy5 tetraodon n
18	2262	32.7	1287	2 Q9VL24 DROME	Q9vl24 drosophila
19	2256	32.7	1287	2 Q9U5W1 DROME	Q9u5w1 drosophila
20	2253	32.6	1287	2 Q7YU59 DROME	Q7yu59 drosophila
21	2205	31.9	1291	2 Q7Q409 ANOGA	Q7q409 anopheles g
22	1864.5	27.0	1223	2 Q9VRC9 DROME	Q9vrc9 drosophila
23	1760	25.5	1003	2 Q7PS03 ANOGA	Q7ps03 anopheles g
24	1725	25.0	1275	2 Q51NK7 MAGGR	Q51nk7 magnaporthe
25	1712.5	24.8	1361	2 Q7XUB7 ORYSA	Q7xub7 oryza sativ
26	1700.5	24.6	1271	2 Q5BBG1 EMENI	Q5bbg1 aspergillus
27	1659.5	24.0	1330	2 Q5KGS9 CRYNE	Q5kgs9 cryptococcu
28	1658.5	24.0	1330	2 Q5S8D4 CRYNE	Q5s8d4 cryptococcu
29	1643	23.8	1273	2 Q4MNG5 ASPFU	Q4mng5 aspergillus
30	1614.5	23.4	1295	2 Q41OK4 GIBZEA	Q41ok4 gibberella
31	1549	22.4	1264	2 Q6BT03 DEBHA	Q6bt03 debaryomyce

RESULT 1

NPCL1 HUMAN					
ID	NPCL1 HUMAN	STANDARD;			
AC	Q9UHC9; Q6R3Q4; Q9UHC8;				
DT	13-SEP-2005 (Rel. 48, Created)				
DT	13-SEP-2005 (Rel. 48, Last sequence update)				
DT	13-SEP-2005 (Rel. 48, Last annotation update)				
DE	Niemann-Pick C1-like protein 1 precursor.				
GN	Name=NPCL1;				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;				
OC	Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2 AND 3), AND TISSUE SPECIFICITY.				
RX	MEDLINE=20247253; PubMed=10783261; DOI=10.1006/geno.2000.6151;				
RA	Davies J.P., Levy B., Ioannou Y.A.;				
RT	"Evidence for a Niemann-Pick C (NPC) gene family: identification and characterization of NPC1L1.";				
RL	Genomics 65:137-145(2000).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), AND TISSUE SPECIFICITY.				
RX	PubMed=14976318; DOI=10.1126/science.1093131;				
RA	Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,				
RA	Tetzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,				
RA	Murgolo N., Graziano M.P.;				
RT	"Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol absorption.";				
RL	Science 303:1201-1204(2004).				
RN	[3]				
RP	TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.				
RX	PubMed=15671032; DOI=10.1074/jbc.M409110200;				
RA	Davies J.P., Scott C., Oishi K., Liapis A., Ioannou Y.A.;				
RT	"Inactivation of NPC1L1 causes multiple lipid transport defects and protects against diet-induced hypercholesterolemia.";				
RL	J. Biol. Chem. 280:12710-12720(2005).				
RN	[4]				
RP	INDUCTION.				
RX	PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;				
RA	van der Veen J.N., Kruit J.K., Havinga R., Baller J.F.W., Chimini G.,				
RA	Lestavel S., Staels B., Groot P.H.E., Groden A.K., Kuipers F.;				
RT	"Reduced cholesterol absorption upon PPARGdelta activation coincides with decreased intestinal expression of NPC1L1.";				
RL	J. Lipid Res. 46:526-534(2005).				
RN	[5]				
RP	FUNCTION.				
RX	PubMed=15928087; DOI=10.1073/pnas.0500269102;				
RA	Garcia-Calvo M., Lissack J., Bull H.G., Hawes B.E., Burnett D.A.,				
RA	Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Detmers P.A.,				
RA	Graziano M.P., Hughes M., MacIntyre D.E., Ogawa A., O'Neill K.A.,				
RA	Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,				
RA	Ujjainwalla F., Altmann S.W., Chapman K.T., Thornberry N.A.;				

Q9shn9 arabidopsis
Q9tvk6 dictyosteli
Q6cbal yarrowia li
Q7rwl9 neurospora
Q59zv0 candida alb
Q55ics dictyosteli
Q4pe83 ustilago ma
Q9svf0 arabidopsis
Q4h344 ciona intes
Q618v4 caenorhabdi
Q19127 caenorhabdi
Q12200 saccharomyc
Q50rb6 entamoeba h
Q750g1 ashbya goss

ALIGNMENTS

Db 301 LLLVGFVAPARDKRMWDPKKGTSLSDKLSFSFTHLLGQFFQGMGTWVASWFLTLVLVS 360
QY 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPNFSGIILDLLELLELQERLRLHQLWSPEAQNRISLQDICVAPLNDPT 480
Db 421 RYDLSLLGPNFSGIILDLLELLELQERLRLHQLWSPEAQNRISLQDICVAPLNDPT 480
QY 481 SYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDMKQFLYCANAPLTKQGTALAL 540
Db 481 SYDCCINSLLQYFQNNRTLLLTANQTLWGQTSQVDMKQFLYCANAPLTKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEBM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEBM 600
QY 601 RAFORMAGMFOVTTAERSLDEINRTTAEPLIPATSYIVIFLYISIALGYSWSRV 660
Db 601 RAFORMAGMFOVTTAERSLDEINRTTAEPLIPATSYIVIFLYISIALGYSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVHIGALGRVAPSMCLCSLSEALCPFLGALTTPMPAVRTTALSGLAV 780
Db 721 YQRLPRRPGEPREVHIGALGRVAPSMCLCSLSEALCPFLGALTTPMPAVRTTALSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPQCGEGLLGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPQCGEGLLGFFQKAYAPPL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFVEGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLPLNRYFVEGAPV 900
QY 901 YFVITLGYNFSSEAGNNAICSSAGCNFSTQKIQVATFPPQSYLAIPASSWVDQFIDW 960
Db 901 YFVITLGYNFSSEAGNNAICSSAGCNFSTQKIQVATFPPQSYLAIPASSWVDQFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNMSITMGSVRPSVEQFHKLPLFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNMSITMGSVRPSVEQFHKLPLFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSNLTSQDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSNLTSQDQVLTVAIPLSRLEYSGTISAHCNLYLLDSASRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPPTITNVFVQVLTTLPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPPTITNVFVQVLTTLPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSVLMLVDVTGPMALWLDISYNAVSILNVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSVLMLVDVTGPMALWLDISYNAVSILNVS 1200
QY 1174 AVGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
QY 1234 LIQIFPFRNLTLTLGLLGLVLPVILSVYVGPVDPNPALEOKRAEVAAMVVASCP 1293
Db 1261 LIQIFPFRNLTLTLGLLGLVLPVILSVYVGPVDPNPALEOKRAEVAAMVVASCP 1320
QY 1294 NHPSRVSTADNLYVNHSPSGISKGAISNPLNNGROF 1332
Db 1321 NHPSRVSTADNLYVNHSPSGISKGAISNPLNNGROF 1359

RESULT 2

NPCL1 RAT
ID NPCL1 RAT STANDARD; PRT; 1331 AA.
AC Q6T3U3;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Niemann-Pick C1-like protein 1 precursor.
GN Name=Npc1l1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND SUBCELLULAR
LOCATION.
RC STRAIN=Sprague-Dawley;
RX PubMed=14976338; DOI=10.1126/science.1093131;
RA Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
Tetzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
Murgolo N., Graziano M.P.,
"Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol
absorption.";
RL Science 303:1201-1204 (2004).
RN [2]
RP SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RX PubMed=1577641; DOI=10.1016/j.bbagen.2004.12.021;
RA Iyer S.P.N., Yao X., Crona J.H., Hoos L.M., Tetzloff G.,
Davis H.R. Jr., Graziano M.P., Altman S.W.,
"Characterization of the putative native and recombinant rat sterol
transporter Niemann-Pick C1 Like 1 (NPCL1) protein.";
RL Biochim. Biophys. Acta 1722:282-292 (2005).
RN [3]
RP FUNCTION.
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
RA Braun C.-P., Crona J.H., Davis H.R. Jr., Dean D.C., Detmers P.A.,
Graziano M.P., Hughes M., Macintyre D.E., Ogawa A., O'Neill K.A.,
Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
Ujjainwalla F., Altman S.W., Chapman K.T., Thornberry N.A.,
"The target of ezetimibe is Niemann-Pick C1-Like 1 (NPCL1).";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137 (2005).
CC -!- FUNCTION: Play a major role in cholesterol homeostasis. Is
critical for the uptake of cholesterol across the plasma membrane
of the intestinal enterocyte. Is the direct molecular target of
ezetimibe, a drug that inhibits cholesterol absorption (By
similarity). Lack of activity leads to multiple lipid transport
defects. The protein may have a function in the transport of
multiple lipids and their homeostasis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
of brush border membranes from proximal enterocytes suggests
considerable association with the apical membrane fraction. Exists
as a predominantly cell surface membrane expressed protein.
CC -!- TISSUE SPECIFICITY: Small intestine showed the highest level of
expression. Expression in other tissue including gall bladder,
liver, testis, and stomach is also observed. Along the duodenum-
ileum axis, the levels vary in different segments of intestine,
with peak expression in the proximal jejunum. Protein expression
is confined to the enterocyte. Discrete localization to the
epithelial layer bordering the luminal space along the crypt-
villus axis. Protein expression in the enterocyte is observed
closest to the luminal space. Expression in enterocytes from the
proximal (jejunum) but not in the distal (ileum) region.
CC -!- PTM: Highly glycosylated.
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.

DR	EMBL; AY437867; AAR97888.1; -, mRNA.	Db	120	FVSLHCHNTCS	PDQSLFINVTRVVERGAGEPPAVVAEAFYQSRPAEKAYESCQVRIPA	179
DR	RGD; I303135; Npc111.	Qy	180	AATLAVGTW	CGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQAVGSIQPLNEG	239
DR	InterPro; IPR004765; NP_C type.	Db	180	AASLAVGSM	CGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPGQALPDGIIQPLNGK	239
DR	InterPro; IPR003392; Patched.	Qy	240	VARCNESQ	GDVATCSCQDCACAPAIARPQALDSTFYLGQMPGSLVLIILCSVFAVVT	299
DR	InterPro; IPR000731; SSD 5TM.	Db	240	IAPCNESQ	GDSDSVCSCQDCACAPVPPPEALRPSFYMGMPGMLAIIIFTAVFVLLS	299
DR	Pfam; PF02460; Patched; I.	Qy	300	ILAVGRVAP	ARDKSMVDPKKTSLSDKLSFSTHLLGQFPQCGWTWVSWPLTILVLS	359
DR	TIGRFAM8; TIGR00917; 2A060601; 1.	Db	300	AVLVRV	VSNRNKNAEGPQAPKLPKHKLSPHTILGRFFQNGWGRVSWPLTIVLALS	359
KW	PROSITE; PS0156; SSD; 1.	Qy	360	VIWVALA	AGLVTELTDPVELWSAPNSQARSEKAFHDQHPFPFRTNVOILTAPNRSS	419
KW	Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;	Db	360	FIWVIALA	AGLTFIELTDPVELWSAPKQARSEKSFHDEHPFPFRTNJOI FVTARNRSS	419
FT	Steroid metabolism; Transmembrane.	Qy	420	YRYDSL	LGLPKNFSGILDLLELLELLELRLHQLWSPEAQRNLSQDICYAPLNPDN	479
FT	SIGNAL	Db	420	YKDSL	LLGSKNFSGILSLDLFLELLELRLHQLWSPEARNISLQDICYAPLNPN	479
FT	CHAIN	Qy	480	TSLVDC	CINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTKDGTALA	539
FT	TOPO DOM	Db	480	TSLDCC	VNSLLOVFQNNRTLLMTANQTLNGQTSQVDKDHFLYCANAPLTKDGTSLA	539
FT	TRANSMEM	Qy	540	LSCNADY	GAPVPEPLAIGGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEE	599
FT	TOPO DOM	Db	540	LSCNADY	GAPVFPFLAVGGYQGTDYSEAEALITFSLNYPADDPRMAQAKLWEEAFLEE	599
FT	TRANSMEM	Qy	600	MRAFQRM	AGMFQVTFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGYSYNSWR	659
FT	TOPO DOM	Db	600	MESQRT	SDKQVAFSAERSLEDEINRTTIQDLPEFAVSYIIVFLYISLALGYSYNSCR	659
FT	TRANSMEM	Qy	660	VMVDSK	ATLGGVAVVLGAVMAANGFSYLGITBSLVLQVWVFLVLSGADNIFVL	719
FT	TOPO DOM	Db	660	VAVESK	ATLGGVIVVLGAVLAAMGFYSYLGVPSSLVIQVVPFLVAVGADNIFIVL	719
FT	TRANSMEM	Qy	720	EYQRL	PRPGPREVHIGRALGRVAPSMLLCSLSEALICFFLGALTPMPAVTFALTSLGA	779
FT	TOPO DOM	Db	720	EYQRL	PRMPGQREAHIGRTLGSVAPSMLLCSLSEALICFFLGALTPMPAVTFALTSLGA	779
FT	CARBOHYD	Qy	780	VILDF	LQMSAFVALLSDSKRQBRASRLDVCCKVKPOBLPPPGQEGELLLGFFQKAYAPF	839
FT	TOPO DOM	Db	780	IILDF	LQMTAFVALLSDSKRQBRASRDVLCFSTRKLPPLPKGKEGILLRFFRKIYAPF	839
FT	CARBOHYD	Qy	840	LLHWT	TRGVLLFLALFGVSLYSNCHTSVGDQELALPKDSYLLDYFLNRYPEVGAP	899
FT	TOPO DOM	Db	840	LLHFR	IRFVVMLLFLTLFGANLYMCNINVGLDQELALPKDSYLLDYFLNRYLEVGP	899
FT	CARBOHYD	Qy	900	VYFVT	TLGVNFSSEAGMNAICSSAGCNFFSQIKIYATEPPEQSYLAIPASSWVDDFD	959
FT	TOPO DOM	Db	900	VYFVT	ISGFNFSSEAGMNAICSSAGCKFSLSLTQKIYASEPDPQSYVAIAASSWVDDFD	959
FT	CARBOHYD	Qy	960	WLTP	-SSCCRLYISGPNKDKFCPSVTNLSNCLNCKMCSITMGSVRSVQSFHYKLPWFLND	1018
FT	TOPO DOM	Db	960	WLTP	SSCCRLYIRGPHKDFCPSDTSFNCLKNCNRTLGPVREPTAEQFHLYLPWFLND	1019
FT	CARBOHYD	Qy	1019	RNIK	CPKGLAAYSTVNLSDGOVLASRFMAVHKPLKNSODYTEALRAARELANITA	1078
FT	TOPO DOM	Db	1020	PNR	ICPKGLAAYSTVNLSDGOVIAFQFMAVHKPLKNSODFTEALRASRLAANITA	1079
FT	CARBOHYD	Qy	1079	DURK	VPDPAPEVFPYITNVFYEQYLTITLPEGLFMLSCLIVPTFAVSCLLGLDLRSG	1138
FT	TOPO DOM	Db	1080	DURK	VPDPAPEVFPYITNVFYEQYLTITLPEGLFMLSCLIVPTFAVSCLLGLDMCSG	1139
FT	CARBOHYD	Qy	1139	LNLS	ISVIMLVDTVGFMALWDISYNAVSLINLVSAGVSVFVSHITSPATKPTWL	1198
FT	TOPO DOM	Db	1140	ILNLS	ISVIMLVDTVGFMALWDISYNAVSLINLVSAGVSVFVSHITSPATKPTWL	1199
FT	CARBOHYD	Qy	1199	ERAK	ATISWGSAGVAVGAVMTNLPGLVLGLAKAQLIQIFFFRNLNLTLLGLLHGLVFL	1258
FT	TOPO DOM	Db	1200	ERAK	ATISWGSAGVAVGAVMTNFPGLLILGLFAQAQLIQIFFFRNLNLTLLGLLHGLVFL	1259

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
Qy 1 MAEAGLGGWLLALLLRLAQSEPTTTHQPGYCAFYDECKNPGLSGMLTSLNVSCLSN 60
Db 1 MAEAGLGGWLLALLLRLAQSEPTTTHQPGYCAFYDECKNPGLSGMLTSLNVSCLSN 59
Qy 61 TPARKITGDHLLILOKICPLRYTGPNTQ-ACCSNKLVSLEASITTKALLTTCPCACSDN 119
Db 60 TPARKITGDHLLILOKICPLRYTGPNTQ-ACCSNKLVSLEASITTKALLTTCPCACSDN 119
Qy 120 FVSLHCHNTCSNPSLEINVTTRVLAQLGAGQLPAVVAEAFYQSRPAEKAYESCQVRIPA 179

QY 1259 PVLSTGVDVNPALALEOKRAEEAAVAVWVWVSCPNHPSRSTADNYYNHSPFSGI-KG 1317
 DB 1260 PVLSTGVDVNPALALEOKRAEEAAVAVWVWVSCPNHPSRSTADNYYNHSPFSGI-KG 1316
 QY 1318 AGAISNLFNNGROF 1332
 DB 1317 ANARSLSPSQKF 1331

RESULT 3

NPCL1_MOUSE
 ID NPCL1_MOUSE STANDARD; PRT; 1333 AA.
 AC G6T3U4; QSSVK1;
 DT 13-SEP-2005 (Rel. 48, Created)
 DT 13-SEP-2005 (Rel. 48, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Niemann-Pick C1-like protein 1 precursor.
 GN Name-Npc1l1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6;
 RX PubMed=14976318; DOI=10.1126/science.1093131;
 RA Altmann S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
 RA Tetzeloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
 RA Murgolo N., Graziano M.P.;
 RA "Niemann-Pick C1 like 1 (NPC1L1) is the intestinal cholesterol
 RT absorption.";
 RL Science 303:1201-1204(2004).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RG The mouse genome sequencing consortium;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP FUNCTION, AND INDUCTION.
 RX PubMed=15173162; DOI=10.1074/jbc.M405817200;
 RA Davis H.R. Jr., Zhu L.-J., Hoos L.M., Tetzeloff G., Maguire M., Liu J.,
 RA Yao X., Iyer S.P.N., Lam M.-H., Lund E.G., Detmers P.A.,
 RA Graziano M.P., Altmann S.W.;
 RA "Niemann-Pick C1 Like 1 (NPC1L1) is the intestinal phytosterol and
 RT cholesterol transporter and a key modulator of whole-body cholesterol
 RT homeostasis.";
 RL J. Biol. Chem. 279:33586-33592(2004).
 RN [4]
 RP TISSUE SPECIFICITY, AND FUNCTION.
 RX PubMed=15671032; DOI=10.1074/jbc.M409110200;
 RA Davies J.P., Scott C., Oishi K., Liapis A., Ioannou Y.A.;
 RA "Inactivation of NPC1L1 causes multiple lipid transport defects and
 RT protects against diet-induced hypercholesterolemia.";
 RL J. Biol. Chem. 280:12710-12720(2005).
 RN [5]
 RP INDUCTION.
 RX PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;
 RA van der Veen J.N., Kruit J.K., Havinga R., Baller J.F.W., Chimini G.,
 RA Lessavel S., Staelen B., Groot P.H.E., Groen A.K., Kuipers F.;
 RA "Reduced cholesterol absorption upon PPARdelta activation coincides
 RT with decreased intestinal expression of NPC1L1.";
 RL J. Lipid Res. 46:526-534(2005).
 CC -!- FUNCTION: Play a major role in cholesterol homeostasis. Is
 CC critical for the uptake of both phytosterol and cholesterol across
 CC the plasma membrane of the intestinal enterocyte. Is the direct
 CC molecular target of ezetimibe, a drug that inhibits cholesterol
 CC absorption (By similarity). Lack of activity leads to multiple
 CC lipid transport defects and exerts a protective effect against
 CC diet-induced hyperlipidemia. Lack of activity also causes a
 CC deregulation of CAV1 transport and localization, suggesting that
 CC the observed lipid transport defect may be the indirect result of
 CC an inability of cells to properly target and/or regulate CAV1

expression. The protein may have a function in the transport of
 multiple lipids and their homeostasis, and may play a critical
 role in regulating lipid metabolism.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
 CC of brush border membranes from proximal enterocytes suggests
 CC considerable association with the apical membrane fraction. Exists
 CC as a predominantly cell surface membrane expressed protein.
 CC -!- TISSUE SPECIFICITY: Expressed in small intestine, stomach and
 CC muscle, along with detectable expression in lung, heart, gall
 CC bladder, brain, testis, skin and liver. Expression in liver is
 CC extremely low.
 CC -!- INDUCTION: Cholesterol/cholelate feeding resulted in down-regulation
 CC of intestinal expression. Expression is decreased by 35% in the
 CC jejunum upon PPARd activation.
 CC -!- PTM: Highly glycosylated (By similarity).
 CC -!- SIMILARITY: Belongs to the patched family.
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AY437866; AAR97887.1; -; mRNA.
 CC EMBL; AL607152; CAI24395.1; -; Genomic DNA.
 CC MGI; MGI:2685089; Npc1l1.
 CC GO; GO:0042632; P:cholesterol homeostasis; IMP.
 CC InterPro; IPR004765; NP_C type.
 CC InterPro; IPR003392; Patched.
 CC InterPro; IPR000731; SSD 5TM.
 CC Pfam; PF02460; Patched; 1.
 CC TIGRFAMs; TIGR00917; 2A060601; 1.
 CC PROSITE; PS50156; SSD; 1.
 CC Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;
 KW Steroid metabolism; Transmembrane.
 FT SIGNAL 1 20 Potential.
 FT CHAIN 21 1333 Niemann-Pick C1-like protein 1.
 FT TOPO_DOM 21 284 Extracellular (Potential).
 FT TRANSMEM 285 305 1 (Potential).
 FT TOPO_DOM 306 352 Cytoplasmic (Potential).
 FT TRANSMEM 353 373 2 (Potential).
 FT TOPO_DOM 374 632 Extracellular (Potential).
 FT TRANSMEM 633 653 3 (Potential).
 FT TOPO_DOM 654 665 Cytoplasmic (Potential).
 FT TRANSMEM 666 686 4 (Potential).
 FT TOPO_DOM 687 696 Extracellular (Potential).
 FT TRANSMEM 697 717 5 (Potential).
 FT TOPO_DOM 718 742 Cytoplasmic (Potential).
 FT TRANSMEM 743 763 6 (Potential).
 FT TOPO_DOM 764 776 Extracellular (Potential).
 FT TRANSMEM 777 797 7 (Potential).
 FT TOPO_DOM 798 846 Cytoplasmic (Potential).
 FT TRANSMEM 847 867 8 (Potential).
 FT TOPO_DOM 868 1113 Extracellular (Potential).
 FT TRANSMEM 1114 1134 9 (Potential).
 FT TOPO_DOM 1135 1142 Cytoplasmic (Potential).
 FT TRANSMEM 1143 1163 10 (Potential).
 FT TOPO_DOM 1164 1165 Extracellular (Potential).
 FT TRANSMEM 1166 1186 11 (Potential).
 FT TOPO_DOM 1187 1206 Cytoplasmic (Potential).
 FT TRANSMEM 1207 1227 12 (Potential).
 FT TOPO_DOM 1228 1242 Extracellular (Potential).
 FT TRANSMEM 1243 1263 13 (Potential).
 FT TOPO_DOM 1264 1333 Cytoplasmic (Potential).
 FT DOMAIN 632 797 SSD.
 FT COMPIAS 267 270 Poly-Pro.
 FT COMPIAS 500 503 Poly-Leu.
 FT CONFLICT 476 476 N -> K (in Ref. 2).
 SQ SEQUENCE 1333 AA; 147132 MW; 7771520D9B352735 CRC64;
 Query Match 78.3%; Score 5407; DB 1; Length 1333;
 Best Local Similarity 77.2%; Pred. No. 0;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
QY	4	AGLRLWALLLLRLAQSEPTTTHOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
DB	3	AAWQWLLWALLLLNSAQELTPTTHKAGCFYFECGKNPELSGGLTSLNSISCLNTPA	62
QY	64	RKITGDHLLILQKICPRLYTGN-TOACCSAKQLVSLASLSITKALLTRCPACSDNFVN	122
DB	63	RHVTGDHLLALLQVCPRLYNGPNDTYACCSTKQLVSLDSSLITKALLTRCPACSENFVS	122
QY	123	LHCHTCSNQSLFNTVTRVAQLGAGOLPAVVAEAFYQHSFASQSDYSCSRVRPAAAT	182
DB	123	IHCNTCSPDQSLFNTVTRVQRPDQLPAPVVAEAFYQHSFASQSDYSCSRVRPAAAS	182
QY	183	LAVGTMCGVYGSAICNAORWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNPGVAR	242
DB	183	LAVGMCGVYGSAICNAORWLNFGQDTGNGLAPLDITFHLLEPGQALADGKPLDGKITP	242
QY	243	CNESQGDVATCSCQDCAAACPAIARPOALDSTFYLGOMPSLVLIILCSVFAVITLL	302
DB	243	CNESQGEDSAACSCQDCAAACSPVIPPFPALRPSFTYMGPMGWLALIIIFTAVFLLSVL	302
QY	303	VGRVAPARDKSMVDPKGTSLSDKLSFSTHTLGGFOGWTGTVASWPLTILVLSVIP	362
DB	303	VYLRVASNNKNTAGSQAPNLPRKRRFSPHTVLGRFESGTVASWPLTVLALSIV	362
QY	363	VVALAAGLVFTLTDPVLSWAPNSQARSEKAFDQHPFPFPRTNQVILTAHNSRYR	422
DB	363	VIALSVGLTFTLTDPVLSWAPNSQARSEKAFDQHPFPFPRTNQVITAKNRSYKY	422
QY	423	DSLILGPKNPSGILDLLELELQERLRHLQVSPQAQRNLSQDICYAPLNPNVSL	482
DB	423	DSLILGPKNPSGILDLLELQELERLRHLQVSPQAQRNLSQDICYAPLNPNVSL	482
QY	483	YDCINSLQYFONNRTLLLTANQTLMTGOTSQVMDKHFLYCANAPLTFKDGITALSC	542
DB	483	TDCVNSLQYFQNNHTLLLTANQTLNGTSLVMDKHFLYCANAPLTYKDGITALSC	542
QY	543	MADYGAPVPFPLAIGGYKGXDEAEALIMTFSNLNYPAGDPRLAQAKUWEAFLEMPA	602
DB	543	IADYGAPVPFPLAVGGVQGTDYSEAEALITFSINNPADPRMAHAKLWEEAFLEMQS	602
QY	603	FORRMAGMFOVTTAERLEDEINRTAEDLPFTATSYIVILYLSLALGSSYSSRVV	662
DB	603	FORSTADKFOIAFSAERLEDEINRTTIQDLPVFAISYLIVLYLSLALGSSYSSRVAV	662
QY	663	DSKATLGLGVAVVLGVAAMGFFSYLGRSLVILQVVPFLVLSVGADNIPFVLEYQ	722
DB	663	DSKATLGLGVAVVLGVAAMGFYSYLVGPSSLVIIQVVPFLVLAAGADNIPFVLEYQ	722
QY	723	RLPRRPGEPREVIHGRALGVAPSMLLCSLSEALCFPLGALTMPAVRTEALTSGLAVIL	782
DB	723	RLPRMPGEQEAHIGRTLGSVAPSMLLCSLSEALCFPLGALTSMPPAVRTEALTSGLAIF	782
QY	783	DFLLQMSAFVALLSLDSKQEAERLDVCCCKQELPPGQSGLLLPFOKAYAPFLH	842
DB	783	DFLLQMTAFVALLSLDSKQEAERPDVVCCFSERNLPPPKQKGLLCLCFRKYITFLH	842
QY	843	WITRGVLLAFALFGVLSYMSCHISVGLDQELAPKDSYLLDYFLFLNRYFVGAPVVF	902
DB	843	RFIRPVVLLFLVLFGANLYLNCNISVGLDQDALPKDSYLYDYFLFLNRYLSEVGPVVF	902
QY	903	VTTILGYNFSSAGNNAICSSAGCNFSFTQIKIYATEFPQSYLAIPASSWDDFDIMLT	962
DB	903	DTTSGYNFSTAGNNAICSSAGCSFSLTKIQYASEFPNQSYVAIAASWDDFDIMLT	962
QY	963	P-SSCCRLYISGPNKDFCPSTVNSLNCNKMSTMGVSVRPSVEQPHKLYLPWFLNDRPN	1021
DB	963	PSSCCRLYITGPHKDFCPSTDTSFNCLKNCMRTLGPVRPTTEQPHKLYLPWFLNDRPN	1022
QY	1022	IKCPKGGLAAYSTSVNLTSQVLAASRPMAHYKPLKNSQDYTEALRAARELANITADLR	1081
DB	1023	IRCPKGGLAAYSTSVNLSSDQIIIASQPMAYHKPLRNSQDPTFALRASRLAANITAE	1082

QY	1082	KVPCTDPAREVPYTTITNVFYQYLTLPGLFMLSCLVPTPAVSCLLGLDRLSGLLN	1141
DB	1083	KVPGTDNPEVFPYTTISNVFYQYLLTVLPEGLFTLALCFVPTFVVCYLLGLDRLSGLLN	1142
QY	1142	LLSIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHITSFALSTKPTWLER	1201
DB	1143	LLSIIMLVDTIGLMAVSGISYNAVSLINLVTAGMSVEFVSHITSFVSTKPTRLERA	1202
QY	1202	KEATISWGSVAFGAVAMTNLPGILVLGLAKAQILQIFFRNLNLLITLGLLHGLVLPVI	1261
DB	1203	KDATTIWMGSAVFAVAMTNPPGILVLGLFAQAQILQIFFRNLNLLITLGLLHGLVLPVV	1262
QY	1262	LSVVGDPVNPALAEOKRAEEAAVAVWVASCNPNRKSVDADNIYVNHSPGS-IKGAGA	1320
DB	1263	LSVLGGDPVNOALVLEEKATEA-AMVSEFSCQPPFPADANTSDYVNYGNPFIFEINA	1321
QY	1321	ISNPLNNGRQF 1332	
DB	1322	ASSSLPKSDQKF 1333	
RESULT 4			
ID	Q4T749	TETNG PRELIMINARY; PRT; 1132 AA.	
AC	Q4T749		
DT	13-SEP-2005	(TrEMBLrel. 31, Created)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last sequence update)	
DT	13-SEP-2005	(TrEMBLrel. 31, Last annotation update)	
DE	Chromosome undetermined SCAF8317, whole genome shotgun sequence.		
DE	(Fragment).		
GN	ORFNames=GSTENG0005930001;		
OS	Tetraodon nigroviridis (Green puffer).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;		
OC	Tetraodontidae; Tetraodontidae; Tetraodon.		
OX	NCBI_TaxID=99983;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RA	Jailion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,		
RA	Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,		
RA	Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,		
RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,		
RA	Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,		
RA	Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,		
RA	Cruaud C., Duprat S., Bottier P., Coutanceau J.P., Gouzy J.,		
RA	Parra G., Lardier G., Chapple C., McKernan K.J., Mesirov J.,		
RA	Kellis M., Volff J.N., Guigo R., Zody M.C., Kahn D., Robinson-Rechavi M.,		
RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn W., Scarpelli C.,		
RA	Laudet V., Schachter V., Quetier P., Saurin W., Scarpelli C.,		
RA	Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.,		
RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals the early vertebrate proto-karyotype."		
RL	Nature 431:946-957(2004).		
RN	[2]		
RG	NUCLEOTIDE SEQUENCE.		
RG	Genoscope; Whitehead Institute Centre for Genome Research;		
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.		
CC	-!- CAUTION: The sequence shown here is derived from an		
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is		
CC	preliminary data.		
DR	EMBL; CAAE01008317; CAF91283.1; -; Genomic_DNA.		
FT	NON TER 1		
SQ	SEQUENCE 1132 AA; 125923 MW; 4283CF4CA0BDE97 CRC64;		

Query Match 46.0%; Score 3179; DB 2; Length 1132;
 Best Local Similarity 52.0%; Pred. NO. 4.3e-216;
 Matches 641; Conservative 167; Mismatches 264; Indels 160; Gaps 11;
 QY 87 TOACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSNQSLFINVTRVAQL- 145
 DB 1 TYACCSINQLSSLEASLASKAVLIRCPSCADNFAHFHCITTCSPQQTITKTKVTNVT 60

Db 4 RGPAGLILLLLCPVQFSQ-----SCWVGECG---IASGD-----KRYNCRYSPPKPL 51

Qy 67 TGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCH 126

Db 52 PEDGYDLVQELCPGFFG-NVSLCCDVQQLRTLDKDLQLPLQFLSRCPSCFFNLMLFCE 110

Qy 127 NTCSPNOSLFINTVR-----VAQLGAGQLPAAVYEAFFQHSFASQSDSCSRVRPAAAT 182

Db 111 LTCSPROQFLNVTATEDYDVPVNTQNTKTNVKELEYVYGETPANAMYNACRDVEAPSSNE 170

Qy 183 LAVGTMCGVYGSAICNAQRWLNFGDDTNGGLAPLDIT--FHLLEPGQAVGSGIQPLNEG 240

Db 171 KALGILLCGREQA-CNATNWTIEMFNKONGQAPFTITPISDLPT-----HGMPEMNAT 224

Qy 241 ARCNSQGDVATCSQDCAASC-----PAIARQALDSTFYLGQMPGSLVLIIL 291

Db 225 KGCDESDEVGTGPCQDCSIVCGPKPQPPPPVPMRILGLDAMVIMWSSYMAFLIYVP 284

Qy 292 CSVFAVTVILLVGRVAPARDKSKMDPKGTSLSKLSFSTHTL----- 336

Db 285 GAFFAVWCY-----RKRYFVSEYTPIDGNIAFVNSSDKGQAFCCDPLGAAF 331

Qy 337 ---LQGFQCGWGTWASMLTILVLSVIPVVALAAGLVFTELTTDPVLMWAPNSQARSE 393

Db 332 ERGLRLFAQMGACVURHPGCVVFFSLAFIVACSSGLVIRVITDPVLMWAPNSQARRE 391

Qy 394 KAFDQHFQPFRTNQVILTAPNRSYRYDSLLGPK-NFSGILDLDLLELLELQERLR 452

Db 392 KEYFDTHFGPFPMQLIIRATNNGSHLYHPVAGADVPPGPPLSRDLHQVLDLQTAIE 451

Qy 453 HLQWSPAQRNISLODICYAPLNPNDTSLVDCINSLLQVFNQNRITLLLTANQTLMGQ 512

Db 452 N--ITASYNNETVTLQDCLAPLSYNK---NCTILSVLNYFQNSHSLV----- 495

Qy 513 TSQV-----DKDHPLYCANAPLTCKDGTALASCWADYGAPVPPFLAIGVKGKDY 564

Db 496 DHQGDFFVYADYHTHFLYCVRAPASLNDASLLHDCPLGTGGVFPVFWLVGVDQDQY 555

Qy 565 SEAEALIMTSLNNYPAGDPRLAQAQKWEAEFLERMAFORRMAGMFOVTTAERSLDE 624

Db 556 NNATALVITFPNNYNTDKLQRAQWSEFINFVKYKN---PNLTISFWAERSIDE 612

Qy 625 INRTAEDLPATSYIVIFLYISALGSSWSRWVDSKATLGLGVAVVVLGVAAM 684

Db 613 LNRESNLDLTILSYAIFMFLYISIALGHKSCSRLVDSKISLGIAGILVLSVACSL 672

Qy 685 GFESVIGIRSSVILQVVPBELVSGADNIRTFVLEYQRLPRRCPGEPREVHIGRALGVA 744

Db 673 GIFSYIGVPLTLVIEVFPFLVLAAGVDNIIFLVQTYQORDERLOQETLDQQLGRVLGVA 732

Qy 745 PSMILCSLSEAI CFFLGALTMPAVRTFALTSGLAVILDFLQMSAFVALLSLDSKQEA 804

Db 733 PSMFLSSSEVAVPLGLSLVVPVHTFSLPAGMAVLIDFLQITCFVSLGLDLKQEK 792

Qy 805 SRLDVCCCKPQELPPPGQ-----EGLLGFQKAYAPFLAHWITRGVWVLLFLALFGVS 860

Db 793 NRLDVVCCVQGAEE---DGAGVQASBSCLPFRFKNSYAPLLKDWMPVIAVAVFGVLSFS 849

Qy 861 LYSMCHISVGLDQELALPKDSVLLDYLFLNRYFVGAPVYVTVTLGYNPSEAGMAIC 920

Db 850 IAVLNKVIIGLDQSLMSMDSDSYMDYFQSLGRYLHAGPPVYFVVBEGHNYTSLKQNNVC 909

Qy 921 SSAGCNFSFTQKIYATFEPPEQSYLAIPAGSWDDFDLWLTTP--SSCCRLYISGNKDKF 979

Db 910 GGLGNNDSLVOQIFTAQLDNYTIFGAPASWIDDYEDNWKPSQSCCRVNS---TDQF 966

Qy 980 CPSTVNSLNCNKMISITM--GSVRPSVBQPHKYLPMFLNDRPNIKCPKGGLAAYSTS 1038

Db 967 CNASVVDPTCIR-CRPLTSEKORQGEDFMFLPMFLSDNPNPKCGKGHAAYSSAVNI 1025

Qy 1039 TSDQO-VLASRPMYHKLKNSQDYTEALRAARELANITADLRKVPDTPAFEVFPYTI 1097

Db 1026 LKNGSGVGATYFMTYHTVLOASADPIDAMQARLIASNIT---RTMGLGESSYRFPYYSV 1082

Qy 1098 TNVFEQYLTITLPEGLFMLSCLVPTPAVSCLLGLDLRSLGILLNLSIVMLVDVTGPM 1157

Db 1083 FVVFYQYLTVIDDTIFNLGVSGLGAIPLAVTVVLMGCELWATVIMCVTIAMILVNMFGVM 1142

Qy 1158 LWDISYNAVSLINLVSAGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVAGVA 1217

Db 1143 LMGISLNAVSLINLVMSCGISVSEFCSHITRAFTLSTKGSRDVRAEALAHMGSSVFSGIT 1202

Qy 1218 MTNLPGILVLGLAKAQLIQIFFFRLNLLITLGLLHGLVPLPVILSVYVGPDPNPALAL 1275

Db 1203 LTKFGGIVLAFAPKASQIFQIFFRMYLAIVLGLGTHGLIFLPLVLSYIGPSINKAKSL 1260

RESULT 7

Q8MKD8_FELCA PRELIMINARY; PRT; 1276 AA.

AC Q8MKD8;

DT 01-OCT-2002 (Tremblrel. 22, Created)

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Niemann-Pick C1.

OS Felis silvestris catus (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;

OC Felis.

OX NCBI_TaxID=9685;

RI [1]

RP NUCLEOTIDE SEQUENCE

RX MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;

RA Somers K.D., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A., Thrall M.A.;

RT "Mutation analysis of feline Niemann-Pick C1 disease.";

RL Mol. Genet. Metab. 79:99-103(2003).

DR EMBL; AF503634; AA27451.1; -; mRNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008158; P:hedgehog receptor activity; IEA.

DR GO; GO:0030301; P:cholesterol transport; IEA.

DR InterPro; IPR004765; NP_C type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD_5TM.

DR Pfam; PF02460; Patched; 1.

DR TIGRFAMs; TIGR00917; 2A060601; 1.

DR PROSITE; PS0156; SSD; 1.

DR SEQUENCE 1276 AA; 141710 MW; C2DD3496472B71EP CRC64;

Query Match 34.8%; Score 2405; DB 2; Length 1276;

Best Local Similarity 40.7%; Pred. No. 3.8e-161;

Matches 535; Conservative 232; Mismatches 461; Indels 86; Gaps 29;

Qy 7 RGMILWALLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCH 126

Db 4 RGPALGLILLLLC---PAQVLAQS---CIWVGECG---IASGD-----KRYNCRYSPPKPL 51

Qy 67 TGDHLILLOKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCH 126

Db 52 PDGYDLVQELCPGFFG-DNVSLLCDVQQLQTLKDLQLPLQFLSRCPSCFFNLVNLFCR 110

Qy 127 NTCSPNOSLFINTVR-----VAQLGAGQLPAAVYEAFFQHSFASQSDSCSRVRPAAAT 182

Db 111 LTCSPROQFLNVTATEDYDVPVNTQNTKTNVKELEYVYGETPANAMYNACRDVEAPSSND 170

Qy 183 LAVGTMCGVYGSAICNAQRWLNFGDDTNGGLAPLDIT--FHLLEPGQAVGSGIQPLNEG 240

Db 171 KALGILLCGKDAE-CNATNWTIEMFNKONGQAPFTITPISDLPT-----HGMPEMNAT 224

Qy 241 ARCNSQGDVATCSQDCAASC-----PAIARQALDSTFYLGQMPGSLVLIIL 291

Db 225 KGCDESDEVGTGPCQDCSIVCGPKPQPPPPVPMRILGLDAMVIMWITMAFLIYVFF 284

Qy 292 CSVFA-----VVTILLVGRVA---PARDKSKMDVPKKGTSLSKLSFSTHTLLGQ 339

Db	285	GAFFALMCKYRKRYFVSEYTPIDSNIAFVSNDRGE-----ASCCDALGAAPGECCLR	337
Qy	340	FFQGGTGWASWPLTILVSVIPVVALAAGLVFTELTPDVELWSAPNSQARSEKAFHQ	399
Db	338	LFSQWGSFCVRNPGPIFFSLAFTACSSGLVFRVTNPNVDLWSAPSSQARLEKEYFDT	397
Qy	400	HFGPFRTNQVILTAPNRSSRYDSLGLPK-NFSGILDLDLLELLELERLHLQWS	458
Db	398	HFGPFRTNQVILTAPNRSSRYDSLGLPK-NFSGILDLDLLELLELERLHLQWS	455
Qy	459	PEAQRNLSQDICVAPLNPNTSYLDCINSLQYFQNNRTLLLTANQITLMGQTSQV--	516
Db	456	SYNNTETVLODICVAPLNPNTSYLDCINSLQYFQNNRTLLLTANQITLMGQTSQV--	507
Qy	517	DKWDFHLYCANAPLTFKDTALALSCMADYCAPVFPFLAIGYKGDYSEAEALIMTFSL	576
Db	508	DYTHLLCYVRAPASLNTSLHDFCLGTGPGVFWLVLGGYDDQNNYNATALVITFPV	567
Qy	577	NNYPAGDPRLAQAKLEAFLEEMAFQRMAGMFOVTTAERSLEDBINRTTAEDLPF	636
Db	568	NNYNDTERLOKAWKEKFNFKVNYKN---PNLTISFTTERTSEDELNESNGDPTV	624
Qy	637	ATSYVIVELYISLALGSSSRSMVDSKATIGLGVAVVLGAVMAAGFFSYGLIRSSL	696
Db	625	IISYAIMELYISLALGSSSRSMVDSKATIGLGVAVVLGAVMAAGFFSYGLIRSSL	684
Qy	697	VILQVVPVLVSGADNIFIFVLEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLBAI	756
Db	685	IIVIEVIFVLAVGVNDIFILVQYQDERLHGETLDQQLGRVLGEVAPSMFLSFSBAV	744
Qy	757	CFPLGALTMPAVRTFALTSGLAVILDFLQMSAFVALLSLDSKQBASRLDVCCKVQ	816
Db	745	AFPLGALSMPAVHTFSLPAGMAVLIDFLQITCFVSLGLDIXKQEKRLDVLCCVRGS	804
Qy	817	ELPPPCQ-CEGILLGFFQKAYAPFLHMTTRGVWLLFLALRGVSLYSNCHISVGLDQEL	875
Db	805	EDGTSVQASESLFRFLKHYSPLLLKQWRPVIATFVGVLSFVAVLNKVEIGLDQSL	864
Qy	876	ALPKDLYLDYFLNRYFEVGPVYFVTTLGVNFSSEAGMNAICSSAGCNFFSTQIKI	935
Db	865	SMDDSYVMDYFKSL-KYLHAGPPVYFVLEEGHDYTSLKGNVCGMGCNDSLVQQIF	923
Qy	936	YATEFPEQSVLAIPASWDDDFDWTLP-SSCRLYISGNKDKPCSTVNSLCKNCKM	994
Db	924	NAAQLDSTYRIGFAPSWDIDYFDWYKQSSCCRVYNS---TDRFCNASVDPACIR-CR	979
Qy	995	SITM-GSVRPSVQEHKYLPMFLNDRPNIKCPKGLAAYSTVN-LTSDGQVLASRFMAY	1052
Db	980	PLTQEGKQRQGGDFWRFLPMFLSDNPNKCGKGHAAYSSAVNLLGNDTGVGATYFMTY	1039
Qy	1053	HKPLKNSQDYTEALRAARELAANITADLRKVPQTDPA-FEVFPYTTNTVFOYLILPE	1111
Db	1040	HTVLQTSADFTDAMRKANLIASNIT---KTMGLEGSNRYRVPYFVYFYEQYLTIID	1095
Qy	1112	GLFMLSCLVPTFVAVSCLLGLDLRSGLLNLSLVMILVDVTCFMAWLDISYNAVSLNL	1171
Db	1096	TIFNLSVSLGAIPLVTVILGCDLWSAVTMCITAMILVMFNMFGVMMWLGISLNAVSLNL	1155
Qy	1172	VSAGMSVZFVSHITRSFALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVILGLAK	1231
Db	1156	VMSCGISVFCSHITRAFTVSMKGSRAQRAEALAHMSSVFGSITLTKPGIWLWAFAK	1215
Qy	1232	AQLQIIFFRNLNLTLLGLHLGLVFLPVILSVGPDVNPALAI---EQRASE 1282	
Db	1216	SOIQIFVFRMYLANVLGATHGLIFLVLVLSYIGPSINKAKSLATQEQYKGTG 1269	
RESULT 8			
Q9GK52 CANFA PRELIMINARY; PRF; 1276 AA.			
AC Q9GK52;			
DT 01-MAR-2001 (Tremblrel. 16, Created)			

DT	01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE	Niemann-Pick type C1 disease protein.
GN	Name=NPCI;
OS	Canis familiaris (Dog).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
OC	Canis
OX	NCBI_TaxID=9615;
RP	[1]
NP	NUCLEOTIDE SEQUENCE.
RA	Murakawa M., Freeman M.W.;
RL	Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AF315034; AG40873.1; -; mRNA.
DR	Ensembl: ENSCARG0000018183; Canis familiaris.
DR	GO: GO:0016021; C: integral to membrane; IEA.
DR	GO: GO:0005764; C: lysosome; IEA.
DR	GO: GO:0016020; C: membrane; IEA.
DR	GO: GO:0008158; F: hedgehog receptor activity; IEA.
DR	GO: GO:0030301; P: cholesterol transport; IEA.
DR	InterPro: IPR004765; NP C type.
DR	InterPro: IPR003392; Patched.
DR	InterPro: IPR000731; SSD_5TM.
DR	Pfam: PF02460; Patched; i.
DR	TIGRFAMs: TIGR00917; 2A060601; 1.
DR	PROSITE: PS50156; SSD; 1.
SQ	SEQUENCE 1276 AA; 141695 MW; 07D1327B51BD5681 CRC64;
Query Match 34.8%; Score 2404.5; DB 2; Length 1276;	
Best Local Similarity 40.4%; Pred. No. 4.2e-161; Indels 105; Gaps 28;	
Matches 524; Conservative 233; Mismatches 434;	

Qy	33	CAFYDEGCKNPELSLMTLSNVSCLSNTPARKITGDHLILQKICPLRYTGNPTACCS	92
Db	25	CVWYEGCG---IASGD---KRYNCQYSGPKPLKDGDLMLQELCPGLFF-DNVS	76
Qy	93	AKQVLSLEASITKALLTRCPACSNFVNLHCHTSCPNQSLFINVTR----	148
Db	77	VQQLRTLKDSLQLPLQLSRPCSFYNLMNLFCELTCSPRQSFNLTETEDYDVPV	136
Qy	149	QLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATAVLAVTMCVYGSALCAQRM	208
Db	137	TKTNVKEQLYVGSFANAMYNACRDVEAPSSNDKALGLLCKEAEA-CNATN	195
Qy	209	TGNGLAPI-----DITFHLEPGQAVGSGIQPLNEGVARCNESQGDVATC	261
Db	196	KONGQAPTIITPISDLPAH-----GKEMNNATKGCDEPVDVETAPCQ	245
Qy	262	SC-----PAIARPOLDSTFVLGQMPGSLVLIILCSVFAVVTILLVGF	312
Db	246	VCGPKPQPPAPAPWRILGLDAMYIMWITTYNAFLMFFGAPFVWCY-----	293
Qy	313	KSKMVDPKGYSLSKLSFSTHT-----LIGQPFQ-----WGTWASW	354
Db	294	-RRYFVSEYTPIDSNIAFVSNAGTGEASCCDALGAAPGECRLRLFTQWGS	352
Qy	355	IIVLSVIVPVALAAGLVFTELTPDVELWSAPNSQARSEKAFHQPFRTNQVIL	414
Db	353	IIFSLAFIAACSSGLVFRVTNPNVDLWSAPSSQARLEKEYFDAHFGPFRT	412
Qy	415	PNRSYRVDLSLLGPK-NFSGILDLDLLELLELERLHLQVSPQAQRNLSQD	473
Db	413	PHTSVHTYQYPSGSDVPFGPPLDIGILHQVLDTQTAIEN---ITATYNN	470
Qy	474	PLNPNTSYLDCINSLQYFQNNRTLLLTANQITLMGQTSQV--DWKDFHLY	531
Db	471	PLSPYK---NCTIMSVNLNFQNSHML-----DHKIGDDFYVADYHLL	522
Qy	532	FKDGTALALSCMADYCAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAG	591
Db	523	LNDTSLHDPCLGTGPGVFWLVLGGYDDQNNYNATALVITFPVNNYNDTE	582
Qy	592	WEEAFLEEMRAFQRRMAGMFOVTTAERSLEDBINRTTAEDLPISATSVI	651

Db 583 WEKEFINFKVYEN---PNTTISFTTERRSIEDLNRESNGDVFVTLISYAVMFLYSIAL 639
QY 652 GSYSSRVWVDSKATLGGVAVVLGAVMAAGRPYSYLGIRSSVLQVLPVPLVLSVCA 711
Db 640 GHKSCSRFLVDSKISLGIAGILVLSWMSGLGIFSGIPUTLIVIEVIFPLVAVG 699
QY 712 DNIFIVLBYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSAICFFLGAITPMPAVRT 771
Db 700 DNIFILVQYQRDERLQGETLEQQLGRVLGEVAPSMFLSSFEAVAPFLGALSQMPAVHT 759
QY 772 FALTSLAVILDFLOMSAFVALLSDSKQERASRLDVCCVKPQELPPGQ---EGL 827
Db 760 FSLFAGMAYLIDFLQITCFVSLGLDLKRQENRDLVLCCLTGS---GGTGIQASESC 816
QY 828 LIGCFKOKAYAPFLHWHITGVVLLFLALFGVLSYMSCHISVGLDELALPKDSYLLDYF 887
Db 817 LRFFFKNSYSPFLKDWMPVIANVFGLISFIAVLNKVEIGLDQSLMPPDSYMDYF 876
QY 888 LFLNRYFEVGAPYFVITTLGYNFSSEAGNNAICSSAGCANNFSTQKIQYATFPPEQSYLA 947
Db 877 KSL-KYLHAGPPVYFVLEBHDVTSLEGQNMVCGMGCCNDLSVQIIPSAQOLDNYTHIG 935
QY 948 IPASSWDFIDWLTIP-SSCRLYISGPNKDFPCSTVNSLCLNKMCSITM-GSVRPSV 1005
Db 936 FAPSSWIDYDFWVKPQSSCCRVTNS---TDQFCNASVVDPAVR-CRPLTQEGKRRPQ 991
QY 1006 EOPHYKLPWFLDRPNKCPKGLAAYSVNLIT-SDGOVLASRMAVHKPLKNSQDYTE 1064
Db 992 EDPFMPFLPFLSDNPNPKCGKGAAGVAVNLVNDTSVGGATYFMTYHVLQTSADFID 1051
QY 1065 ALRAARELAANTADILRKVPDPAPEVPPYITNTVFFYEQYLITLPEGLFMLSCLVPTF 1124
Db 1052 AMRKAILIASNITKM-SLKGS-VRVPYSVFFVYEQYLITLIDDTIFNLVSGLATF 1108
QY 1125 AVSCLLGLDLRSGLNLNLISVILVDTVFGMALMDISNAVSLNLINLAVSGMSVEFVSH 1184
Db 1109 LVTLLVGLCELASVIMCTVITAMILVMPGVMLWGLISLNAVSLNVLNMSCGISVEFCSH 1168
QY 1185 ITRSPAISNPTWLERAKATISMSGAVFAGVAMTNLPGILVLGLAKALQIIFPRLNL 1244
Db 1169 ITRAFTVSANGRVERKAEALSHMSSVFSGLITLTKFGGIVVLAFAKSQIFQVITGMVL 1228
QY 1245 LITLLGLHLVFLPVILSYGPDVNPALAL-EQKR 1279
Db 1229 ANVLGATHGLFLPVLISYIGPSINKAKSLASQER 1264
AC O15178; Q9P130; STANDARD; PRT; 1278 AA.
DT 30-MAY-2000 (Rel. 39, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Niemann-Pick C1 protein precursor.
GN Name=NPCL1
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE, VARIANT ILE-642, AND VARIANTS NPCL1.
RX MEDLINE=97362323; PubMed=9211849; DOI=10.1126/science.277.5323.228;
RA Carstea E.D., Morris J.A., Coleman K.G., Loftus S.K., Zhang D.,
RA Cummings C., Gu J., Rosenfeld M.A., Pavan W.J., Krizman D.B.,
RA Nagle J., Polymeropoulos M.H., Sturley S.L., Ioannou Y.A.,
RA Higgins M.E., Comly M., Cooney A., Brown A., Kaneski C.R.,
RA Blanchette-Mackie E.J., Dwyer N.K., Neufeld E.B., Chang T.-Y.,
RA Liscum L., Strauss J.F. III, Ohno K., Zeigler M., Carmi R., Sokol J.,
RA Markie D., O'Neill R.R., van Diggelen O.P., Ellender M.,
RA Patterson M.C., Brady R.O., Vanier M.T., Pentchev P.G., Tagle D.A.;

RT "Niemann-Pick C1 disease gene: homology to mediators of cholesterol
RT homeostasis.";
RL Science 277:228-231 (1997).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND VARIANTS.
RX MEDLINE=99355599; PubMed=10425213; DOI=10.1006/bbr.1999.1070;
RA Morris J.A., Zhang D., Coleman K.G., Nagle J., Pentchev P.G.,
RA Carstea E.D.;
RT "The genomic organization and polymorphism analysis of the human
RT Niemann-Pick C1 gene.";
RL Biochem. Biophys. Res. Commun. 261:493-498 (1999).
RN [3]
RP NUCLEOTIDE SEQUENCE, AND VARIANT ILE-642.
RX MEDLINE=21623216; PubMed=11754101; DOI=10.1002/humu.10016;
RA Bauer P., Knoblich R., Bauer C., Finckh U., Hufen A., Kropp J.,
RA Braun S., Kustermann-Kuhn B., Schmidt D., Harzer K., Rolfs A.;
RT "NPC1: complete genomic sequence, mutation analysis, and
RT characterization of haplotypes.";
RL Hum. Mutat. 19:30-38 (2002).
RN [4]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA], AND VARIANT ILE-642.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.J., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywicki M.I., Skaleka U., Smallos D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=99128318; PubMed=9927649; DOI=10.1073/pnas.96.3.805;
RA Watari H., Blanchette-Mackie E.J., Dwyer N.K., Glick J.M., Patel S.,
RA Neufeld E.B., Brady R.O., Pentchev P.G., Strauss J.F. III;
RT "Niemann-Pick C1 protein: obligatory roles for N-terminal domains and
RT lysosomal targeting in cholesterol mobilization.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:805-810 (1999).
RN [6]
RP VARIANT NPD TRP-992.
RX MEDLINE=9829797; PubMed=9634529;
RA Greer W.L., Riddell D.C., Gillan T.L., Girouard G.S., Sparrow S.M.,
RA Byers D.M., Dobson M.J., Neumann P.E.;
RT "The Nova Scotia (type D) form of Niemann-Pick disease is caused by a
RT G3097-->T transversion in NPC1.";
RL Am. J. Hum. Genet. 63:52-54 (1998).
RN [7]
RP VARIANTS NPC1 GLN-934; LEU-940; ASN-948; TRP-992; ALA-1007;
RP THR-1061 AND VAL-1213.
RX MEDLINE=99452586; PubMed=10521290;
RA Greer W.L., Dobson M.J., Girouard G.S., Byers D.M., Riddell D.C.,
RA Neumann P.E.;
RT "Mutations in NPC1 highlight a conserved NPC1-specific cysteine-rich
RT domain.";
RL Am. J. Hum. Genet. 65:1252-1260 (1999).
RN [8]
RP VARIANT NPCL1 THR-1061.
RX MEDLINE=99452593; PubMed=10521297;
RA Millat G., Marcals C., Rafi M.A., Yamamoto T., Morris J.A.,
RA Pentchev P.G., Ohno K., Wenger D.A., Vanier M.T.;

"Niemann-Pick C1 disease: the I1061T substitution is a frequent mutant allele in patients of Western European descent and correlates with a classic juvenile phenotype.";
Am. J. Hum. Genet. 65:1321-1329(1999).
[9]
RN VARIANTS NPC1, AND VARIANTS ARG-215; VAL-858 AND GIN-1266.
RX MEDLINE=99408226; PubMed=10480349; DOI=10.1007/s004390051057; Zhang H., Yamamoto T., Namba E., Nimomiya H., Higaki K., Taniguchi M., Akaboshi S., Watanabe Y., Takeshima T., Inui K., Okada S., Tanaka A., Sakuragawa N., Millat G., Vanier M.T., Morris J.A., Pentchev P.G., Ohno K.;
RA "NPC1 gene mutations in Japanese patients with Niemann-Pick disease type C.";
RT Hum. Genet. 105:10-16(1999).
[10]
RN VARIANTS NPC1 GLN-958 AND ALA-1007.
RX MEDLINE=2113111; PubMed=11349231;
RA Sun X., Marks D.L., Park W.D., Wheatley C.L., Puri V., O'Brien J.F., Kratt D.L., Lundquist P.A., Patterson M.C., Pagano R.E., Snow K.;
RA "Niemann-Pick C variant detection by altered sphingolipid trafficking and correlation with mutations within a specific domain of NPC1.";
RT Am. J. Hum. Genet. 68:1361-1372(2001).
[11]
RN VARIANTS NPC1 ALA-378; MET-950; ARG-992 AND THR-1061.
RX MEDLINE=21313105; PubMed=11333381;
RA Millat G., Marçais C., Tomasetto C., Chikh K., Fensom A.H., Harzer K., Wenger D.A., Ohno K., Vanier M.T.;
RA "Niemann-Pick C1 disease: correlations between NPC1 mutations, levels of NPC1 protein, and phenotypes emphasize the functional significance of the putative sterol-sensing domain and of the cysteine-rich luminal loop.";
RT Am. J. Hum. Genet. 68:1373-1385(2001).
[12]
RN VARIANTS NPC1 TYR-177; CYS-978 AND VAL-1035.
RX MEDLINE=21372069; PubMed=11479732; DOI=10.1007/s004390100531;
RA Ribeiro I., Marcao A., Amaral O., Sa Miranda M.C., Vanier M.T., Millat G.;
RA "Niemann-Pick type C disease: NPC1 mutations associated with severe and mild cellular cholesterol trafficking alterations.";
RT Hum. Genet. 109:24-32(2001).
CC -I- FUNCTION: Involved in the intracellular trafficking of cholesterol. May play a role in vesicular trafficking in glia, a process that may be crucial for maintaining the structural and functional integrity of nerve terminals.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Found in late endosomes and lysosomes.
CC -I- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain containing a di-leucine motif necessary for lysosomal targeting are critical for mobilization of cholesterol from lysosomes.
CC -I- PTM: Glycosylated.
CC -I- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease type C1 (NPC1) [MIM:257220]. NPC1 is an autosomal recessive lipid storage disorder, which affects particularly the brain, liver and spleen, and which is characterized by lysosomal accumulation of low density lipoprotein derived cholesterol. Clinical features include variable hepatosplenomegaly and severe progressive neurological dysfunction such as ataxia, dystonia and dementia. The age of onset can vary from infancy to late adulthood.
CC -I- DISEASE: Defects in NPC1 are the cause of Niemann-Pick disease type D (NPD) [MIM:257220]; also known as Niemann-Pick disease without sphingomyelinase deficiency, or Nova Scotian type. Because of evidence from biochemical changes, lack of complementation, and linkage mapping to the same chromosome site, NPD and NPC1 are considered to be allelic disorders.
CC -I- SIMILARITY: Belongs to the patched family.
CC -I- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

DR EMBL; AF002020; AAB63982.1; -; mRNA.
DR EMBL; AF157379; AAD48006.1; -; Genomic DNA.
DR EMBL; AF157365; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157366; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157367; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157368; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157369; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157370; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157371; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157372; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157373; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157374; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157375; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157376; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157377; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157378; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF157379; AAD48006.1; JOINED; Genomic DNA.
DR EMBL; AF123046; AAF28875.1; -; Genomic DNA.
DR EMBL; AF123045; AAF28875.1; JOINED; Genomic DNA.
DR EMBL; BC063302; AAF63302.1; -; mRNA.
DR EMBL; ENSG00000141458; Homo sapiens.
DR HGNC; HGNC:7897; NPC1.
DR MIM; 607623; -.
DR MIM; 257220; -.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0005764; C:lysosome; TAS.
DR GO; GO:0005624; C:membrane fraction; TAS.
DR GO; GO:0005478; F:intracellular transporter activity; TAS.
DR GO; GO:0015248; F:sterol transporter activity; TAS.
DR GO; GO:0004888; F:transmembrane receptor activity; TAS.
DR InterPro; IPR004765; NP_C_type.

Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 5.8e-161; Gaps 25;
Matches 522; Conservative 237; Mismatches 450; Indels 103;

QY 7 RGLWLLALLRLAQSEPTTTHQFCYCAFVDECG-----KNPELGSGLMTLSNVSLCS 59
DB 4 RGLALGLLLLLLCPAQVFSG-----SCWYGEGLAYGDKRYNCEYS----- 46
QY 60 NTPARKITGDHLLLOKICPLRYTGPNTQACCSAKQVLSLEASISITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG--NVSLCCDVRQLQTLKDNQLQPLQFLSRCPSCFN 103
QY 120 FVNLCHNTCSPNOSLFINVTR---VAQLGAGQLPAVYAEAFYQHSFASQSDSCSRV 175
DB 104 LLNLFCELTCSPRQSOFNLVATEDYDPVNTQTKNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RYPAATILAVGTMCVYGSALCNQORLNFQDGTGNGLAPLDT-----PHLLEPGA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWTYEMFNKDNQGOAPFTITPVFSDPPVH----- 215
QY 229 VCSGQPLNEGVARCNESQGDVATCSQDCAASCAPATARPQ-----ALDSTFYIG 279
DB 216 ---GHEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILILCSVFVAVVTILL-----VGRVAPARPKSMQVKKGTSLIS 326
DB 273 WITYWAFLLVFGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFGQGTWVASWPLTLVLVSIVPVVALAGLVETELTTPDVELWSAP 386
DB 325 DPVSAAFEGCLRLFLTRWGSFCVRNPGCVIFFSFLFITACSSGLVFRVVTTPVDLWSAP 384
QY 387 NQCARSEKAFHDFHGFPRFTNQVILLTAPNRSSYRYSLLLGPK-NFSGILDLLLELL 445
DB 385 SSOARLEKBYFDQHFPGFPFTTQLIRAPLTQKHLYQPYPSGADVPFGPLDIQLHQLV 444
QY 446 ELQERLRHQLQVMSPEAQRNLSQDICYAPLNPDTSLYDCCINSLLQYQNNRLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLLQDLCIAPLSVPNT---NCTILSVLNYFQNSHVLHKK 499
QY 506 NQTLMGQTSQVDWKDHFLYCANAPLTFKDTALALSCMADYGAQVFPFLAIGYKGDYS 565

Db 500 GDDFF---VYADYHTHFLYCVAPASLNDTSLHDFCLGTGCGPVPFWLVLGGYDDQNTN 556
 QY EAEALIMTSLNNYPAGDPRLAQAKLWEBAFLMEAFQRMAGMPOVFTFAERSLEDEI 625
 Db 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISFTAERSLEDEI 613
 QY 626 NRTAEDLPFATSYIVIFLYISLALGSSYSSMRVMDVSKATLGLGVAVVILGAVMAWG 685
 Db 614 NRESDSVFTWISYAIMFLYISLALGHMKSCRLLVDSGLGIAGILVSSVACSIG 673
 QY 686 PFSYLGISSVLQVVPVLVSVGADNIFIVLEYQRLPRPCEPREVHIGRALGRVAP 745
 Db 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQORDERLQGETLDQOLGRVLEAVP 733
 QY 746 SMLLCSLSEAI CFFGALTPMPAVRTFALTSGVLAVLDFLLQWSAPVALISDSKQRAS 805
 Db 734 SMLFSFSFTVAFPLGALSVMRAVHTFSLPAGLAVFIDELQITCFVSLGLDIKQKKN 793
 QY 806 RLDDCCVQKQBLPPPGQ--GEGLLGFQKAVAPFLLHWTITRGVLLFLALFGVSLYSM 864
 Db 794 RLDFICVCGAEDGTSVQASESCLFRFFKNYSPLLLKDMWRPIVIAIFVGVLSPSIAVL 853
 QY 865 CHISVGLDQELPKDSYLLDFLPLNRYFEVGPVYFTYTLGYNPSSSEAGNAICSSAG 924
 Db 854 NKVDIGLQSLMPDSDSYMDVYFKSISQYLHAGPPVYFVLEBHDYTSKQGNMVGCGMG 913
 QY 925 CNFSPFQIKQVATEPEQSYLAIPASSWVDDPIDMLTP--SSCCRLYISGPNKDKPCPT 983
 Db 914 CNDSLVQOIFNAQDLNTRTGFAPSSMIDDYFVWKQSSCCRV---DNITDQPCNAS 970
 QY 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYTSVN--LTS 1040
 Db 971 VVDPACVR--CRPLTBGKQRPQGGDFMRFLPFLSDNPNPKCGKGGHAYSSAVNILLGH 1029
 QY 1041 DQVLAARMVAKPLKNSQDTEALRAARELAANITADLRKVPDTPAFVFPYITIVN 1100
 Db 1030 GTRVGATYEMTYHTVLTQTSADFIDALKKARLIASNVT--ETMGINGS--AYRVFPYSVFV 1086
 QY 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGFMALWD 1160
 Db 1087 FTEQVLTIDDTIFNLGVSIGAIFLVTMVLGCELSAVIMCATIAMVLVNMFGVWMLMG 1146
 QY 1161 ISYNAVSLNLVSAVMSVEFVSHITRSPAI STKPTWLERAKEATISMGSAVPAVAMTN 1220
 Db 1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVSMKGSVERAEBALAHMSSVFSGITLTK 1206
 QY 1221 LFGILVLGLAKAQLOIQIFPRNLTLTLGLHLGVFLPVLVILSYVGPDVNPA 1272
 Db 1207 FGGIVVLAPAKSQIQIIFYFRMYLAMVLIGATHGLIFLPLVLSYIGPSYNKA 1258

RESULT 10
 Q59GR1 HUMAN :
 ID Q59GR1 HUMAN PRELIMINARY; PRT; 1289 AA.
 AC Q59GR1;
 DT 10-MAY-2005 (TrEMBLrel. 30, Created)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
 DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
 DE Niemann-Pick disease, type C1 variant (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Brain;
 RA Totoki Y., Toyoda A., Takeda T., Sakaki Y., Tanaka A., Yokoyama S.,
 RA Ohara O., Nagase T., Kikuno F.R.;
 RT "None title."
 RL Submitted (MAR-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB209048; BAD92285.1; -; -.

FT NON TER 1 1
 SQ SEQUENCE 1289 AA; 143230 MW; 663DB27C13FD801P CRC64;
 Query Match 34.8%; Score 2402; DB 2; Length 1289;
 Best Local Similarity 39.7%; Pred.No. 6.3e-161;
 Matches 525; Conservative 239; Mismatches 452; Indels 106; Gaps 26;
 QY 2 ABAGLRGMLLWA-----LLRLAQSBEPYTHIOPGYCAFYDECG-----KNPELSG 47
 Db 3 ASAALRGHSMRTARGALGALLLCPAQVPSQ-----SCVMYGECCGIAYGDKRYNCEYSG 57
 QY 48 SLMTLSNVSCLSNTPARKITGDHLLILQKICRLTYTGPNTQACCSAKQLVSLSEASITK 107
 Db 58 -----PPKPLPKDGYDLVQLBCLCPFFFG-NVSLCCDVRQLQTLKDNLQLPL 102
 QY 108 ALLTRCPACSDNFVNLHCHNTCSPNOSLFINVTR---VAQLGAGOLPAVVAEAPYQHS 163
 Db 103 QFLSKPCSPCYMLLHLCBELTCSQRSQPLVNTATEDYVDPVNTQTKTNKBLQYTVGGS 162
 QY 164 PAEQSYDSCSRVVRPAAATLAVGTMCVGYGSALCNAQRWLNFGDGTNGGLAPLDTIT--- 219
 Db 163 PANAMYNACRDVEAFSSNDKALGLLGGKADA--CNATNWIEMENKDNQAPFTITPVFS 221
 QY 220 -PHLLPQAVSGIGLOPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO----- 270
 Db 222 DFPV-----RGMEPMNATKGCDESDEVTA PCSCQDCSIVCGPKQPPPPAPWTI 273
 QY 271 -ALDSTFVLGQMGPSLVLILLCVSFAVVTILL-----VGRVAPARDKSKM 316
 Db 274 LGLDAMVIMTYMAFLVFPFGAFVAVCYRKRYPVSEYTPIDSNIAFSV--NASDKGB- 331
 QY 317 VDPKKGTSLSLDSFTHTLQGFQGMGTWASWPLTILVLSVLPVVAALAGLVTELT 376
 Db 332 -----ASCCDPVSAAPEGCLRELTRWGSFCVRNPGCVIFSLVFIACSSGLVEVRT 385
 QY 377 TDPVELMSPNSQASSEKAFHQHFGPFRFTNQVILTAPNRSYRYDSLLGLPK--NFSGI 435
 Db 386 TNPVDLMSAPSQARLEKEYFDQHFQFPFRTQLIIRAPLTDKHIYQYPSPGADVFPGGP 445
 QY 436 LDLDLLELLELLEQLRLHLQWSPEAQRNISLQDICYAPLNPNTSLYDCCINSLLQYEQ 495
 Db 446 LDQILHQLVLDLQIAIEN---ITASVDNETVTLQDLCPLSPYNT---NCTLSVLNLYFQ 500
 QY 496 NNRTLLLTANTLGMQTSQVDWKDHFLYCANAPITFKDGTALALS CMADYGA VPFPFLA 555
 Db 501 NSHSVLDHKKGDDFP---VYADYHTHFLYCVAPASLNDTSLHDFCLGTGCGPVPFWL 557
 QY 556 IGGYKGDYSEABALIMTFSLNYPAGDPRLAQAKLWEBAFLMEAFQRMAGMPOVFT 615
 Db 558 LGGYDDQNTNATLAVITFPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTISF 614
 QY 616 TAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSMRVMDVSKATLGLGVAV 675
 Db 615 TAERSLEDELNRESDSVFTWISYAIMFLYISLALGHMKSCRLLVDSGLGIAGILV 674
 QY 676 VLGA VMAAMGPFYSYLGISSVLQVVPVLVSVGADNIFIVLEYQRLPRPCEPREVH 735
 Db 675 VLSVACSLGVSPSYGLPLTLVIEVIFPLVAVGVDNIFILVQAYQORDERLQGETLDQ 734
 QY 736 IGRALGRVAPSMLLCSLSEAI CFFGALTPMPAVRTFALTSGVLAVLDFLLQWSAPVAL 795
 Db 735 LGRVLGEVAPSMFLSFSFSETVAFPLGALSVMRAVHTFSLPAGLAVFIDELQITCFVSL 794
 QY 796 SLDSPQEAASRLDVCCKVPQELPPGQ--GEGLLGFQKAVAPFLLHWTITRGVLLFL 854
 Db 795 GLDIKQKKNRLDIFCCVRGAEDGTSVQASESCLFRFFKNYSPLLLKDMWRPIVIAIFV 854
 QY 855 ALFGVSLYSVMCHISVGLDQELPKDSYLLDFLPLNRYFEVGPVYFTYTLGYNPSSSE 914
 Db 855 GVLSPSIAVLNKNVDIGLQSLMPDSDSYMDVYFKSISQYLHAGPPVYFVLEBHDYTSK 914
 QY 915 GNNATCSSAGCNPFSFTQIKQVATEPEQSYLAIPASSWVDDPIDMLTP--SSCCRLYISG 973

Db 915 GQMVCGGCGNNDLSVOOIFNAQOLDNYTRIGFAPSSWIDDYDFWVKPQSSCCRV---D 971

Qy 974 PNKDFCFSTVNSLNCNCKMSIT-MGSVRPSVEQFKYLPWFNDNRPNIKPCPKGGLAAY 1032

Db 972 NITDQFCNVSVDPAQVR-CRLPTGKQRPQGGDFMFLPWFSLSDNPNKCGKGHAAY 1030

Qy 1033 STSVN--LTSQGVLASPMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAP 1090

Db 1031 SSANILLGHGTRVGATVMTYHTVLOTSDAFIDALKKARLTASVNT-ETWGIINGS--AY 1087

Qy 1091 EYFVPTITNVFEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLLLSVMILV 1150

Db 1088 RVFVSVFVFEQVLTIDITIFNLGVSGLGAFILVTWVLGCELSAVIMCATIAMVLV 1147

Qy 1151 DTVGFMALWDISYNAVSLINLSAVGMSVBFVSHITRSPASTKPTWLERAKEATISMG 1210

Db 1148 NMFVGMWLGISINAVSLVNLVMSGIVSEFCSHITRAFTVSMKGSVERAEALAHMS 1207

Qy 1211 AVFAGVAMVNLPGIILVGLAKAQLQIOPFRNLNLTILGLLHGLVFLPVLSYVGPVN 1270

Db 1208 SVFSGITLTKFGGIVVLAFAKSQIQIFYFRMYLAMVLLGATHGLIFLFPVLSYIGPSV 1267

Qy 1271 PA 1272

Db 1268 KA 1269

RESULT 11

Q7TMD4_MOUSE

ID Q7TMD4_MOUSE PRELIMINARY; PRT; 1277 AA.

AC Q7TMD4; 25, Created

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)

DE Niemann Pick type C1.

GN Name=Npcl;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP NUCLEOTIDE SEQUENCE

RC STRAIN=C57BL/6; TISSUE=Brain;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butlerfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length human

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.

RN [3]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=C57BL/6; TISSUE=Brain;

RA Strausberg R.;

RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC052437; AH52437.1; -; mRNA.

DR EMBL; BC054539; AH54539.1; -; mRNA.

DR GO; GO:0006897; P: endocytosis; IMP.

DR InterPro; IPR004765; NP_C type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD 5TM.

DR Pfam; PF02460; Patched_1.

DR TIGRFAMs; TIGR00911; ZAO60601; 1.

DR PROSITE; PS0156; SSD; 1.

SQ SEQUENCE 1277 AA; 142882 MW; 3B42230AAC8E564E CRC64;

Query Match 34.7%; Score 2400; DB 2; Length 1277;

Best Local Similarity 40.4%; Pred. No. 8.7e-161;

Matches 528; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

Qy 11 LMAILLRLAOSPEYTHIQPGYCAFYDECGNPELSGLMTLSNVSLNTPARKITGDH 70

Db 8 LGLLLLLCPAQVFSQ-----SCVWYGEQG---IATGD---KRYNCKYSGPPKPLPDG 55

Qy 71 LILQKICPLRYTQNTQACCSAKQLVSLASISITKALLTRCPACSDNFVNLHCHMTCS 130

Db 56 YDLVQELCPGLFP-DNVSLLCCDIOQLTKENLQPLQFLSRCSCFYNLMTLECELTCS 114

Qy 131 PNQSLFINVTRVAQLGAGQLPA----VVAYEAFYQHSFAQSQSDYSCSRVRVPAATLAVG 186

Db 115 PHQQLNVTATEDYFDPKTKQENTNKELEYFVGQSFANAMYNACRDVEAPSSNEKALG 174

Qy 187 TMCVGYSGALCNAORWLNFGQDTGNGLAPLDI-----TFHLLPEQQAQVSGIQLNEGVA 241

Db 175 LLCGRDARA-CNATNWIEMFNKNGQAPFTIIPVFSDSLIL-----GMEPMRNATK 225

Qy 242 RCNCSQGDVATCSQDCQCAASC-----PATARPOALDSTF-----YLGQMPG 283

Db 226 GCVESVDEVTGPCSCQDCSIVCGPKPQPPPPMPKRWINGLDAMYIMVTVYVFLFVFFG 285

Qy 284 SILVILILCSVFVAVVITLLVGRFVAPARDKSKMVDPKGTSLSKLSKLSSTHTLLGQFQG 343

Db 286 ALLAVWCHRRYFVSEYTPIDSNIAFVNS--DKGEASCCDPLGAFFDDCLRMFTK 341

Qy 344 WGTWASWPLTILVLSVTPVVALAAGLYFTETLTPVELWSAPNSQASERKAFHDPFGP 403

Db 342 WGAFCVRNPTCIIFESLAFITVCSGLFVQVTTNPVELWSAPNSQASERKAFHDPFGP 401

Qy 404 PRTNQVILTAPNRSSYRYSLLGLPK-NFSGILDLLELLELLELLELLELLELLELLELLE 462

Db 402 PRTTQLIIQANTSVHLYEPYPAGADVPFGPPLNKEILHQVLDLQ-----IAIESITAS 456

Qy 463 RN---ISLQDICYAPLNPTSLYDCCINSLLQYFQNRRTLLLTANQTLMGQTSQV--- 516

Db 457 YNNETVTLQDICIAPLSPYK---NCTIMSVLNYPQNSHAVL-----DSQVGD 502

Qy 517 -----DWKDHFLYCANAPLTFFDGTALALSCHADYGAPEVFLAIGGYKGDYSEAEALI 571

Db 503 FYIYADYHFLYCYVRAPASLNDTSLHGPCUGTGGPVPFWLVGGYDDQNNATLV 562

Qy 572 MTFSLNNYPAGPRLAQAKLWEAEFLERNAPQRRMAGMFQVTTAERSLEDEINRTTAE 631

Db 563 IFFPVNNYNDYERLQRAWAKEFISFVKYKN---PNLTISFTAESRTEDBELNESIS 619

Qy 632 DLPFATSYIVFLYISLALGSYSWSRVWVDSKATLGLGGVAVVLGAVWAAAGFFSYLG 691

Db 620 DVFTVIISYVMFLYISLALGHQSCSRLNVDSKISLGIAGILVIVLSSVACSLGIFSYMG 679

Qy 692 IRSSLVILQVPPVLVLSVGADNIFIVLEYQRLPRPGEPRVHIGRAGLRVAPSMLLCS 751

Db 680 MPLTLIVIEVLPFLVAVGVDNIFILVQYQDRERLQETLDQQLGRILGEVAPTMFLSS 739

Qy 752 LSEAIFFFLGALTPMPAVRTPALTSGILAVILFLQMSAFVALLSDSKQESRLDVCC 811

Db 740 FSETSAFFFGALSSMPAVHTFSLFAGMAVLIDFLQITCFVSLGLGDIKQEKHLDILC 799

812	Qy	CYKQBELPPGQG-----EGLLGFQKAYAPFLLHWITRGVULLLFLALFGVSLYSMCHI	867
800	Db	CVRGAD---DQCGSHASESLYFRFKNYFAPLLKWLRLPIVAVFVGVLSFSVAVNVKV	856
868	Qy	SVGLDQELALPKQSYLDDYFLFNRYPEVGAPVVFVTTLGVNFSSEAGMNAICSSAGCINN	927
857	Db	DIGLDQSLSPENDSVYIDYFKSLAQYLHSGPPVYFVLEEGYVNSRRKGQNMVCGMGCDN	916
928	Qy	FSFTQKIQYATEPPEQSYLAIPASWVDDDFIDMLTP--SSCCRLYISGNKDKFCPSTVNS	986
917	Db	DSLVOQJFNAEALDYYTRGVGFAPSSWDDYDFDMVSPQSSCCRLY---NVTHQFCNASVMD	973
987	Qy	LNCLKNQWSTIT-MGSVRPSVEQHFHYLPWFINDRPNIKCPKGLGAAYSTSVNLTSQGQVL	1045
974	Db	PTCVR-CRPHTPGKQBPQSGFKMFLPMFLSDNPNPKCKGKGAAYGSAVINVGDDTYI	1032
1046	Qy	ASRPMAYHKLKNSQDYTEALRAARELAANITADLKVPDGTDPAFEPVFPYTTITNVFYEQ	1104
1033	Db	GATVFMYHTLTKTSADYTDAMKKARLIASNITETWBS-KGSD--YRVFPVSVFYVPYEQ	1089
1105	Qy	YLITLPGELFWLSCLVPTFAVSCLLIGLDIRSGLLNLSIVMLIVDTVTCFMAINDISYN	1164
1090	Db	YLTIIIDTITFNLVSLSGIEFLVTLVGLCELWSAVINCIITAMLTVMNFMGVMLWLGISLN	1149
1165	Qy	AVSLINLVSAGMSVEFVSHTTSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGI	1224
1150	Db	AVSLVNLVMSGISVEFCSHITRAFTMSTKGSVRSAEALAHMGSSVFSGITLTKFGGI	1209
1225	Qy	LVLGLAKAQIIOIPFFPRNLNLTLLGLLHGLVFLPVILSVTGVGDVNDPA	1272
1210	Db	VWLAFAKSOFEIPFYPMYLAWVLLGATHGLIIFLVPVLSYIGSVNKA	1257

RESULT 12

RESUDI 12
09N000 FELCA

Q2000 TECH
ID 09N000 FELCA PRELIMINARY: PRT: 1276 AA.

AC 09N000:

DT 01-OCT-2000 (TREMBlrel. 15. Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Niemann-Pick type C1 disease protein.

GN Name=NPCL;

OS *Felis silvestris catus* (Cat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;

OC Felis.

OX NCBI_TaxID=9685;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC TISSUE=Kidney;

RA Murakawa M., Freeman M.W.;

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF258783; AAF72187.1; -; mRNA.

DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0005764; C:lysosome; IEA.

DR GU; GU:0016020; C:membrane; IKA.

DR; GO:0008158; F:hedgenog receptor activity
DB; GO:0003031; P:cholesterol transport; IPA

DR GO; GO:0030301; F:cholesterol transport; IEA.
DR Interpro: IPR004765; NP C type

DR InceIfio; IPR004765; NF_C type.
DR InterPro: IPR003392; patched.

DR IntelPro: IPR000731: SSD 5TM-
DR IntelPro: IPR003392: Facched.

DR PFam: PF02460: Patched: 1.
DR InChIKey: IFR000731; SSD 3M.
DR

DR TIGRFAMS: TIGR00917: 2A050601

DR PROSITE; PS50156; SSD; 1.

SQ	SEQUENCE	1276 AA;	141739 MW;

[illegible]

Query Match

Best Local Similarity 40.6%; Pred. No. 1e-160;

Matches 534; Conservative 232; Mismatches 46

7 RGWLLWALLRLA0SEPYTTIHOPGYCAFYDECCKNPGLSGSLMTLSNVCLSNTPARKI 66

[illegible]

Qy	1112	GLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGMALWDISYNAVSLNL	1171
Db	1096	TIFNLSVSLGAIFLVTVLLGCDLWSAVIMCIITAMILVNMFGVWMLWGLSLNAVSLVNL	1155
Qy	1172	VSANCMSEVFVSHITRSFAITKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAK	1231
Db	1156	VMSCGISVEFCSHITRAFTVSMKSGRAQRAEALAHMGSSVFSGITLTKFGGIVVLAPAK	1215
Qy	1232	AQLIQIFFRNLMLITLGLLHGLVFLFVILSVGPDVNPALAL---EOKRAE	1282
Db	1216	SOIFQIFYFRMYLANVLLGATHGLIFLPLVLLSYIGPSINKAKSLATQOYKGTGTE	1269
RESULT 13			
Q8M149_FELCA			
ID	Q8M149	FELCA PRELIMINARY; PRT; 1276 AA.	
AC	Q8M149		
DT	01-OCT-2002	(Tremblrel. 22, Created)	
DT	01-OCT-2002	(Tremblrel. 22, Last sequence update)	
DT	01-OCT-2003	(Tremblrel. 25, Last annotation update)	
DE	Mutant Niemann-Pick C1.		
OS	Felis silvestris catus (Cat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;		
OC	Felis.		
OX	NCBI_TaxID=9685;		
RP	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RX	MEDLINE=22695580; PubMed=12809639; DOI=10.1016/S1096-7192(03)00074-X;		
RA	Somers K.L., Royals M.A., Carstea E.D., Rafi M.A., Wenger D.A.,		
RA	Thrall M.A.;		
RT	"Mutation analysis of feline Niemann-Pick C1 disease.";		
RL	Mol. Genet. Metab. 79:99-103(2003).		
DR	EMBL; AF503633; AAM27450.1; -; mRNA.		
DR	GO: GO:0016021; C:integral to membrane; IEA.		
DR	GO: GO:0005764; C:lysosome; IEA.		
DR	GO: GO:0016020; C:membrane; IEA.		
DR	GO: GO:0008158; F:hedgehog receptor activity; IEA.		
DR	GO: GO:0030301; P:cholesterol transport; IEA.		
DR	InterPro; IPR004765; NP_C type.		
DR	InterPro; IPR003392; Patched.		
DR	Pfam; PF02460; Patched: I.		
DR	TIGRFAMs; TIGR00917; 2A060601; 1.		
DR	PROSITE; PS0156; SSD; 1.		
SQ	SEQUENCE 1276 AA; 141694 MW; CB3562214195998B CRC64;		
Query Match 34.7%; Score 2395; DB 2; Length 1276;			
Best Local Similarity 40.6%; Pred. No. 2e-160;			
Matches 534; Conservative 232; Mismatches 462; Indels 86; Gaps 29;			
Qy	7	RGMLWALLRLAQSBEYTHQGYCAFYDECGKNPELSGLMTLSNVCSLNTNPARKI	66
Db	4	RGPALGELLLLC---PAQLAQS--CIWYGECC---IASGD---KRYNCKYSGPKPL	51
Qy	67	TGDLHLLLOKICPLRYTGNTQACCSAKQLVLSLEASISITKALLTRCPACSDNFVNLHCH	126
Db	52	PKQGYDLVQLCPGFF--DNVSLCCDVQQLTKDNLQLPLQLSLRCPSCFYNLNLFCE	110
Qy	127	NTCSPNQSLFINVTR---VAQLGAGQLPAVAYEAFYQHSFABQSQYDSCSRVRVPAAT	182
Db	111	LTCSPRQSQFLNVATATEDYDVPVNTQKTNVKGELQYYIGESFANAMYNACRDVEAPSSND	170
Qy	183	LAVGTMCVGVGSAICNAQRMNLFQDGTGNGLAPLDIT--PHLEPGQAVSGIQLPNEG	240
Db	171	KALGLLCKGDAEA--CNATNMIEYMFSGDKNGQAPFTTIPFSDLPFT-----HGM	224
Qy	241	ARCNESQGDVATCSQDCDCAASC-----PAIARPDALDSTFYLGMFGSLVLIIL	291
Db	225	KGCDSEVDVETGPCSQDCSIVCGPKQPPPPVPPWRILGLDAMVIMTWIYMAFLVFP	284
Qy	292	CSVFA-----VVTILLVGRVA---PARDKSKMDPKKGTSLSDKLSFSTHTLLGQ	339

Db	285	GAPFALWCYKRYFVSEYTPIDSNIAFVSNDRGE-----ASCCDALGAAPFECCLRR	337
Qy	340	FFCGWGTWASWPLTTLVLVSVPVVALAAGLVFTELTTPVLEWLSAPNSQARKEKAFHQ	399
Db	338	LFSQWGSFCVRNPGPIFFSLAFIACSSGLVFVRVITNPVDSWLSAPSSQARLEKEYFDT	397
Qy	400	HFGPFRTNQVILTAPNRSSRYRDSLLGLPK-NFSGILDLDLLELLEQLERHLHQVMS	458
Db	398	HFGPFRTTEQLIIQAHTSAHTYQYPGSDVFPGLDLAILHQVLDLQTAIEN--ITA	455
Qy	459	PEAQRNISODICVAPLNDPNTSLYDCCINSLLOYFQNNRTLLLTANQTLMGQTSOV--	516
Db	456	SYNNETVTLQDICVAPLSPYNK---NCTILSVLNIFQNSHML-----DHEIGDDFFVYA	507
Qy	517	DKKDHELYCANAPLTCKDGTALALSCWADYGAVPFPLAIGGYKDYKDYSAEALIMTFSL	576
Db	508	DYTHLLYCVRAPASLNDTSLLDHDPCLGTGPGFPVPLVLGGYDDQNNYNATALVITFPV	567
Qy	577	NNYPAGDPRLAQAQKLEAEFLEEMRAFORMAGHFQVTFTAERSLEDEINRTTAEDLPIF	636
Db	568	NNYNDTERLOKAHWEKEFINFVQYKN---PNLTISFTTERTSIEDELNRESNGDFTV	624
Qy	637	ATSVIVFLYISALGYSWSRVMDSKATLGLGVAVVLGAVMAAMGPFYSIGIRSSL	696
Db	625	IISYAIMFLYISTALGHKSCSRLLDVSKISLGIAGLILVLSVACSIGLIFSIVGIPLTL	684
Qy	697	VILQVPPFLVLSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAI	756
Db	685	IVIEVIFPLVLAQVDNIFILVQTYQDERLHGETLQQLGRVLGEVAPSMFLSPSEAV	744
Qy	757	CFFLGALTPPAVTRFALTGLAVILDLLQMSAFVALLSDSKRQASRLDVCCKVQP	816
Db	745	AFFLGALSKMPAVHTFSLFAGMAVLIDFLQITCFVSLGLDIDIKRQEKRLDVLCCVRGS	804
Qy	817	ELPPPGQ-GEGLLGGFPQKAYAPELLHWITGVVLLFLALFGVLSVSMCHISVGLDQEL	875
Db	805	EDGTSVQASECLPRLFKHSYSPLLDKWMPVIAIFVGVLSFSAVLNKEVIGLDQSL	864
Qy	876	ALPKDSYLLDYFLNRYFEVGAPVYFVTTILGYNFSSSEAGMNAICSSAGCYNFSTQIKI	935
Db	865	SNPDDSYVMDYFKSL-KYLHAGPPVYFVLEBHDYVTSLKGNQMVCGGMCNDSLVQQIF	923
Qy	936	YATEPPEOSYLAIPASSWDDFDWLP-SSCCHLYISGPNKDKFCPSVNSLNLKNCM	994
Db	924	NAAQDSYTRIGFAPSSWIDDYFDWVKPQSSSCRYNS---TDRFCNASVVDPAICR-CR	979
Qy	995	SITM-GSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN-LTSDGOVLASRFMAY	1052
Db	980	PLTQEGKQRPQGGDFMRFLPMLSDNPNPKCGKGHAAYSSAVNLGNDTGVGATYFTY	1039
Qy	1053	HKPLKNSQDYTEALRAARELAANITADLRKVPQTDPA-FEVPYTYITVTFEYQYLTLPE	1111
Db	1040	HTVLQTSADFTDAMERKANLIASNIT---KTMGLEGSNRYRPPYVSVFYFYEQYLTIID	1095
Qy	1112	GLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGMALWDISYNAVSLNL	1171
Db	1096	TIFNLSVSLGAIFLAVTVLLGCDLWSAIVMCIITAMILVNMFGVWMLWGLSLNAVSLNL	1155
Qy	1172	VSANCMSEVFVSHITRSFAITKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAK	1231
Db	1156	VMSCGISVEFCSHITRAFTVSKSGRAQRAEALAHMGSSVFSGITLTKFGIVVLAPAK	1215
Qy	1232	AQLIQIFFRNLMLITLGLLHGLVFLFVILSVGPDVNPALAL---EQKRAE 1282	
Db	1216	SOIFQIFYFRMYLANVLLGATHGLIFLPLVLLSYIGPSINKAKSLATQBOYKGTGTE 1269	

RESULT 14
NPCL MOUSE
ID NPCL_MOUSE STANDARD; PRT; 1278 AA.
AC O35604; O35605;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)

10-MAY-2005 (Rel. 47, Last annotation update)
 Niemann-Pick C1 protein precursor.
 Mus musculus (Mouse).
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 NUCLEOTIDE SEQUENCE.
 MEDLINE:97362324; PubMed=9211850; DOI=10.1126/science.277.5323.232;
 Loftus S.K., Morris J.A., Carstee E.D., Gu J.Z., Cummings C.,
 Brown A., Ellison J., Ohno K., Rosenfeld M.A., Tagle D.A.,
 Pentchev P.G., Pavan W.J.;
 "Murine model of Niemann-Pick C disease: mutation in a cholesterol
 homeostasis gene.";
 Science 277:232-235(1997).
 [2]
 SUBCELLULAR LOCATION.
 MEDLINE:99145606; PubMed=9990080; DOI=10.1073/pnas.96.4.1657;
 Patel S.C., Suresh S., Kumar U., Hu C.Y., Cooney A.,
 Blanchette-Mackie E.J., Neufeld E.B., Patel R.C., Brady R.O.,
 Patel Y.C., Pentchev P.G., Ong W.-Y.;
 "Localization of Niemann-Pick C1 protein in astrocytes: implications
 for neuronal degeneration in Niemann-Pick type C disease.";
 Proc. Natl. Acad. Sci. U.S.A. 96:1657-1662(1999).
 CC -!- FUNCTION: Involved in the intracellular trafficking of
 cholesterol. May play a role in vesicular trafficking in glia, a
 process that may be crucial for maintaining the structural and
 functional integrity of nerve terminals.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late
 endosomes and lysosomes.
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in perisynaptic
 astrocytic glial processes. Also expressed in heart, spleen, lung,
 liver, skeletal muscle, kidney, testis.
 CC -!- INDUCTION: Activated by the drugs progesterone and U-18666A which
 block cholesterol transport out of lysosomes and by the
 lysosomotropic agent NH4Cl.
 CC -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain
 are critical for mobilization of cholesterol for lysosomal targeting
 CC -!- SIMILARITY: Belongs to the patched family.
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC
 This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC
 EMBL; AF003348; AAB63372.1; -; mRNA.
 CC
 EMBL; AF003349; AAB63373.1; -; Genomic_DNA.
 CC
 PIR; T30188; T30188.
 CC
 Ensembl; ENSMUSG0000024413; Mus musculus.
 CC
 GO; GO:0006897; P: endocytosis; IMP.
 CC
 InterPro; IPR004765; NP_C type.
 CC
 InterPro; IPR003392; Patched.
 CC
 InterPro; IPR000731; SSD_5TM.
 CC
 Pfam; PF02460; Patched; 1.
 CC
 TIGRFAMs; TIGR00917; 2A060601; 1.
 CC
 PROSITE; PS0156; SSD; 1.
 CC
 Glycoprotein; Lysosome; Signal; Transmembrane.
 KW
 SIGNAL 1 23 Potential.
 FT CHAIN 1 23 Potential.
 FT Niemann-Pick C1 protein.
 FT TRANSMEM 271 291 Potential.
 FT TRANSMEM 352 372 Potential.
 FT TRANSMEM 623 643 Potential.
 FT TRANSMEM 655 675 Potential.
 FT TRANSMEM 685 705 Potential.
 FT TRANSMEM 761 781 Potential.
 FT TRANSMEM 834 854 Potential.
 FT TRANSMEM 1099 1119 Potential.

FT	TRANSMEM	1125	1145	Potential.
FT	TRANSMEM	1196	1216	Potential.
FT	TRANSMEM	1228	1248	Potential.
FT	DOMAIN	621	786	SSD.
FT	MOTIF	1275	1278	DI-leucine motif.
FT	COMEBIAS	250	260	Poly-Pro.
FT	CARBOHYD	71	71	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	123	123	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	138	138	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	186	186	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	223	223	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	229	229	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	415	415	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	460	460	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	525	525	N-linked (GlcNAc. .) (Potential).
FT	CONFLICT	474	474	S -> F (in Ref. 1; AAB63373).
FT	CONFLICT	479	479	N -> D (in Ref. 1; AAB63373).
FT	SEQUENCE	1278	AA; 142889	MM; 43C71CE47D283674 CRC64;

Query Match 34.5%; Score 2385; DB 1; Length 1278;
 Best Local Similarity 40.2%; Pred. No. 1e-159;
 Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY	14	LLRLAQSEPYTHIQGYCAFYDECGKNPELSSGLMTLSNVCSLSTNPARKITGDHLIL	73
DB	12	LLLLLCPAQVFSQ-----SCWYEGCG-----IATGD-----KRYNCKYSGPPKPLPKDGYDL	59
QY	74	LQKICRLTYGTNTQACCSAKQIVLSLEASLSITKALLTRCPACSDNFVNLHCINTCSPNQ	133
DB	60	VQELCPGLFF-DNVSLLCCDIQQLQTLKSNLQLPLQFLSRCPSCFYNLMTLFCELTCSPHQ	118
QY	134	SLPINVTRVAQLGAGQLPA---VWAYEAFYQHSFAESYSDCSRSRVVPAATAVAGTWC	189
DB	119	SQFLNVTATEDYDPKTPENKTNVKELEYVQGSFANAMYNACRDVEASSEKALGLLIC	178
QY	190	GVYGSALCNAQRWLNFGQDTGNGLAPLDI-----TFHLEPGQAVSGIGIQLNIEGVARN	244
DB	179	GRDARA-CNATNWIEYFMKNGQAPFTIIPVPSDLSIL-----GMEPMENATKGVN	229
QY	245	ESQGDVATCSQDCNASC-----PAIARPOALDSTP-----YLQMPGSLV	286
DB	230	ESVDEVTCSCQDCSIVCGPKPQPPPPMPWRIMGLDAMYIMWYVYAFVFPFGALL	289
QY	287	LIILCSVPFVAVTLLVGRVPAPADKSKWDPKGTSLSDKLSPTHTLLIGOFFQGWGT	346
DB	290	AWCHRRRYFVSEYTPIDSNIAFSVNS-----DKGEASCCDPLGAADFCLRMFTKGA	345
QY	347	WVASWPLTILVLVSIPWVALAGLVFTLTTPDELWSAPNSQARSEKAFHQHGFPPFR	406
DB	346	PCVRNPTCIIPFSLAFITVCSGLVQVVTNPVELWSAPHQARLEKEYFDKHPGPPFR	405
QY	407	TNQVILTAPNRSSYRDSLLGPK-NFSGILDLDDLELLELLELLELLELLELLELLELLE	464
DB	406	TEQLIQAPNTSVHITYEYPAGADVPFGPLNKEILHQLVNLQ-----IAIESITASVNN	460
QY	465	--ISLQDICVAPLNPDNTSLYDCCNSLIQYFQNNRTLLLTANQLMGQTSQV-----	516
DB	461	ETVTLQDICVAPLSPYK-----NCTIMSVNLVYFQNSHAVL-----DSQVGDFFY	506
QY	517	--DWKDHFLYCANAPLTFKDGFTALALSCMADYGAIVFPFPLAIGYGYKDYSAEALIMTF	574
DB	507	YADYTHFLYCYRAPASLNDTSLHGPCCLGTGGVFPFWLVLGGYDDQNNYNATALVITP	566
QY	575	SLNNYPAGDPRLAQAKLWEBAFLBEMRAPQRMAGHQVTFPTARSLDEINRTTADPLP	634
DB	567	PVNNTYNDTERLQRAWAKEKEPISFVKYKN---ENLITISPTARSIEDELRESNDVP	623
QY	635	IPATSVIVFLYISALQSYSSWSVMDSKATLGLGGVAVVLGVAVMAWAGFYSYLGIRS	694
DB	624	TVIISVVMFLYISALGHIQCSRLVLSKISLIAGILIVLSSVACSLGIFSYNGMPL	683
QY	695	SIVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGPREVHTIGRALGRVAPSMLLCSLSE	754

Db	684	TLIVIEVPEFLVAVGVNIFILVQYQDRERLQEBETIDLOQLGRILGEVAPTMFLSFSE	743	34.3%; Score 2367.5; DB 2; Length 1277;
Qy	755	AICPFLGALTMPAVRTALTSLAVILDFLLQMSAFVALLSDSKQESRLDVCCKV	814	Best Local Similarity 40.1%; Pred. No. 1.7e-158;
Db	744	TSAPFGALSSMPAVHTFSLFAGMAVLDFLLQITCFVSLGLDIDKQEKHLIDLCVR	803	Matches 517; Conservative 227; Mismatches 445; Indels 101; Gaps 26;
Qy	815	POELPPGQ---EGLLGFQKAYAPFLHWTIRGVVLLFLALFGLSVLSYMSCHISVG	870	
Db	804	GAD---DQGSASHSESYLFRFPKYNFAPLLKWLRLPIVAVFVGLVSFAVNVKVDIG	860	
Qy	871	LDQELALPKDSVLLDYFLNRYFEVGAPVTFVTTLGNFNSSEAGMAICSSAGCNESF	930	
Db	861	LDQSLMPNDSVIANFKSLAQYLHSGPPVYFVLEEGYNSRKGQNMVCGMGCDNDSL	920	
Qy	931	TQKIYATEFFQOSYLAIAPASSWDDFDLWLP--SSCCRLYISGPNKDKFCPSTVNSLNC	989	
Db	921	VQOIFNAABELTYTRVGFAPSWIDDYFDWVSFQSSCCRLY---NVTHQFCNASVMDPTC	977	
Qy	990	LKNCMSIT--MGSVRPSVEQFHLYLFWFLNDRENIKCPKGLLAAYSTSVNLTSDGOVL--AS	1047	
Db	978	VR--CRPLTEPGKQKQKQKFMFLPFLSDNPNPKGKGHAAGSAVNIUGDDTYIGAT	1036	
Qy	1048	RFMAYHKLKNSODYTEALRAARELANITADLRKVGTDPAPFVPPYITINVFYEQYLT	1107	
Db	1037	YFMYTHILKTSADYTDAMKARLIASNITETMRS--KGSQ--YRVFPYSVYFVYEQYLT	1093	
Qy	1108	ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWDISYNAVS	1167	
Db	1094	IIDDITFNLSVLSGSLFVTLVVLGCLMSAVIMCTIAMILNVNFGVWMLGSLNAVS	1153	
Qy	1168	LINLVSAGMSVEFVSHITRSFAISTKPTWLERAKETISMGSAFVAGVAMTNLPGILVL	1227	
Db	1154	LYNLVNSCGISVEFCHITRAFTMTKGSVRSAEALAHMSSVPSGITLTKFGGIVVL	1213	
Qy	1228	GLAKAQLIQQFPRMLNLLTLGLLGLVLPVLSYVGPVNPA	1272	
Db	1214	AFAKSQIFEIFYFRMYLAMVLLGATHGLIFLPVLLSYIIGPSVNKA	1258	
RESULT 15				
Qy	Q9GLC9_BOVIN	PRELIMINARY;	PRT;	1277 AA.
Id	Q9GLC9; BOVIN PRELIMINARY; PRT; 1277 AA.			
AC	Q9GLC9;			
DT	01-MAR-2001 (TrEMBLrel. 16, Created)			
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Niemann-Pick type C1 disease protein.			
GN	Name=NPC1;			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;			
OC	Pecora; Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	NUCLEOTIDE SEQUENCE.			
RC	TISSUE=Aorta endothelium;			
RA	Murakawa M., Freeman M.W.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AF299073; AG24620.1; -, mRNA.			
DR	GO; GO:0016021; C:integral to membrane; IEA.			
DR	GO; GO:0005764; C:lysosome; IEA.			
DR	GO; GO:0016020; C:membrane; IEA.			
DR	GO; GO:0008158; F:hedgehog receptor activity; IEA.			
DR	GO; GO:0030301; P:cholesterol transport; IEA.			
DR	InterPro; IPR004765; NP C type.			
DR	InterPro; IPR003392; Patched.			
DR	InterPro; IPR000731; SSD_5TM.			
DR	Pfam; PF02460; Patched; 1.			
DR	TIGRFAMs; TIGR00917; 2A060601; 1.			
DR	PROSITE; PS0156; SSD; 1.			
DR	SEQUENCE 1277 AA; 141911 MW; 30C91ABFDIABFICF CRC64;			
Qy	1007	QFHKYLPWFENDRPNIKCPKGLLAAYSTSVNLTSDG--QVLASRFMAYHKLKNSQDYTEA	1065	
Qy	33	CAFYDCGKNPELSGLMTLSNVCSLNTSPARKITGDHLLQKICPLRYTGNTOACCS	92	
Db	25	CIWYEGC---IASGD---KRYNCRYSGPPEPLPDQGYDLVQELCPGFFG--NVSLCCD	76	
Qy	93	AKQIVSEBASLSTIKALLTRCPACSNFVNLHCHNTCSNQSIFNVTRVAQL---GAGQ	149	
Db	77	VQQLHTLKNLQLOFLQSRCPSCFVNLVNLFCBLTCSPPQSOFNLVNTATEDYVDPATNQ	136	
Qy	150	LPAVAYEAFY--QHSFAEQSYDSCSRVRPAAATLAVGTMCGVYGYSALCMAQRLWPNQD	208	
Db	137	TKTNVRELQYVGESFANAMYNCSRVDPEAPSSNEKALGLLCGREASA--CNATNWIETMFN	195	
Qy	209	TGNGLAPLDTT--PHLLEPGQAVGSGIQIPLNEGVARNCSQGGDDVATCSQDCDCAASCPAI	266	
Db	196	KDNGQAPFTITPVFSDLPT-----HGMPEPMNATKGCDESVDVETGFCSCQDCSVCVCPK	250	
Qy	267	ARPO-----ALDSTFVLGQMPGSLVLIILILCSFAVVVITLLVG--FRVAPARDK--	313	
Db	251	PQPPPPVPWKILGLDAMY-----VIMWSTYMAFLVFFGAFVAVCYRKYF	298	
Qy	314	-----SKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV	360	
Db	299	VSEFTPIDGNIPFSINASDKGGPTCCDPLGAAPENHLRLFEWMSFCVRHPGCVVFFSV	358	
Qy	361	IPVVALAAGLVFTBLTTPVELWSAPNSQARSEKAFHDQHGFPFRPNQVILLTAPNRSY	420	
Db	359	AFIAACSSGLVFIQVTTDPVDLMSAPGQARLEKEYEDTHFGFPFRTEQLIIRAPHTPPH	418	
Qy	421	RYDSLLGLPK--NFGSGLDLDLLELLBLQERLRLHQLVMSPEAQRNLSIQDIICYAPLNPDN	479	
Db	419	IYEPYSGADVPPGPPPLAVNLHQLVLDLQTAIE--SITASVNNETVTLRDICVAPLSPYN	476	
Qy	480	TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQV--DWKHFLICANAPLTFKQGT	537	
Db	477	---QNCITLSVLYNFQNSHVL-----DHQVGDDFFVYADYTHFLYVVRPAVSLNDTSL	528	
Qy	538	LALSCMADYGAIPVPFPFLAIGGYKGYSEAEALIMTSLNNYPAGDPLAOKLWEEAPL	597	
Db	529	LHDPCLGTFGGVPFWLVGDDQNNYNATLVITPPVNNYNDNTEKLOQAQWKEFI	588	
Qy	598	EMRAFORRMAGMFOVTFEASRLDEINRTTAEIDLPIFATSYIVIFLYIISLAGSYSSW	657	
Db	589	NPQNYEN---PNLTISFAERSIEDELNRNSDVFVTLISYGMFLYISIALGHKISC	645	
Qy	658	SRVMVDSKATLGLGCVAVVLGAVMAAMGFFSYLGRSSILVQLVPPFLVSLVAGDNIFIF	717	
Db	646	RELLVDSKILLIAGVILVSPVACSLGFFSYMGSPPLTVIEVIFPLVAVGVNIFIL	705	
Qy	718	VLEYQRLPRPGEPREVIHGRALGRVAPSMLLCSSEALCPFLGALTTPMPAVRTFALTSG	777	
Db	706	VQTYQIDERLOQETLDQGVRLGVEVAFSMFLSSFAETVAFPLGLSLVMPAVHTLSLPA	765	
Qy	778	LAVILDFLLQMSAFVALLSDSKQESRLDVCCKVQFQELPPPG-----QEGGLJ	828	
Db	766	MAVLDFLLQITCFVSLGLDIDKQEKQLDVLCCV-----GGAADDAGTQASESCL	817	
Qy	829	LGFFQKAYAPFLHWTIRGVVLLFLALFGLSVLSYMSCHISVGLDQELAPKDSYLLDYFL	888	
Db	818	FRFRNSYAPLLKDWMRPLVAVFVGLSFIAVLNKEIIGLDOSLSMPPDSYVTDYFQ	877	
Qy	889	FLNRYFEVGAPVYFVTLGYNFSSBAGMAICSSAGCNFNFSPTKIQYATEPEQSLEYAI	948	
Db	878	SLNQYLHAGPPVTVFLEEGHDYTSYKQNMVCGGLGCNNDLSVQQVFTAAQLDSYTRIGF	937	
Qy	949	PASSWDDFDLWLP--SSCCRLYISGPNKDKFCPSTVNSLNCINLKNCSIT--MGSVRPSVE	1006	
Db	938	APSSWIDDYFDWVKPQSSCCRIYNS---TEQFCNASVNVNPTCVR--CRPLTPEGKQPOGA	993	

THIS PAGE IS BLANK

GenCore version 5.1.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:21:19 ; Search time 52 Seconds
(without alignments)
2117.769 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNRGQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
.Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/taa/5 COMB.pep.*
- 2: /cgn2_6/prodata/1/taa/6 COMB.pep.*
- 3: /cgn2_6/prodata/1/taa/H COMB.pep.*
- 4: /cgn2_6/prodata/1/taa/PCTUS COMB.pep.*
- 5: /cgn2_6/prodata/1/taa/RE COMB.pep.*
- 6: /cgn2_6/prodata/1/taa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	34.8	1318	2	US-09-949-016-10152
2	2402.5	34.8	1278	2	US-09-462-136-2
3	2385	34.5	1319	2	US-09-462-136-4
4	1329	19.2	1170	2	US-09-462-136-6
5	1078	15.6	1296	2	US-09-462-136-9
6	651.5	9.4	1447	1	US-08-540-406-19
7	651.5	9.4	1447	2	US-08-656-055-19
8	651.5	9.4	1447	2	US-08-954-668-19
9	651.5	9.4	1447	2	US-09-268-140-5
10	651.5	9.4	1447	2	US-08-918-658-19
11	651.5	9.4	1447	2	US-09-724-631-19
12	651.5	9.4	1447	2	US-08-954-701A-19
13	651.5	9.4	1447	2	US-09-807-007-6
14	651.5	9.4	1447	2	US-09-754-032-19
15	651.5	9.4	1447	2	US-08-916-140-19
16	651.5	9.4	1447	4	PCT-US95-13233-19
17	646.5	9.4	1434	1	US-08-540-406-10
18	646.5	9.4	1434	2	US-08-656-055-10
19	646.5	9.4	1434	2	US-08-954-668-10
20	646.5	9.4	1434	2	US-08-918-658-10
21	646.5	9.4	1434	2	US-09-724-631-10
22	646.5	9.4	1434	2	US-08-954-701A-10
23	646.5	9.4	1434	2	US-09-754-032-10
24	646.5	9.4	1434	2	US-08-916-140-10
25	646.5	9.4	1434	4	PCT-US95-13233-10
26	639.5	9.3	642	2	US-09-248-796A-16161
27	628	9.1	1296	2	US-08-857-636-60

28	584	8.5	1203	2	US-09-293-505-2	Sequence 2, Appli
29	584	8.5	1203	2	US-09-060-939A-2	Sequence 2, Appli
30	583.5	8.4	1182	2	US-09-293-505-7	Sequence 7, Appli
31	583.5	8.4	1182	2	US-09-060-939A-7	Sequence 7, Appli
32	581.5	8.4	531	2	US-09-248-796A-16160	Sequence 16160, A
33	580	8.4	1203	2	US-09-207-857-2	Sequence 2, Appli
34	580	8.4	1203	2	US-09-909-280A-2	Sequence 2, Appli
35	579	8.4	1146	2	US-09-807-007-1	Sequence 1, Appli
36	526.5	7.6	1311	1	US-08-540-406-4	Sequence 4, Appli
37	526.5	7.6	1311	2	US-08-656-055-4	Sequence 4, Appli
38	526.5	7.6	1311	2	US-08-954-668-4	Sequence 4, Appli
39	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
40	526.5	7.6	1311	2	US-09-724-631-4	Sequence 4, Appli
41	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
42	526.5	7.6	1311	2	US-09-754-032-4	Sequence 4, Appli
43	526.5	7.6	1311	4	US-08-916-140-4	Sequence 4, Appli
44	526.5	7.6	1311	4	PCT-US95-13233-4	Sequence 4, Appli
45	474	6.9	1286	2	US-09-268-140-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-949-016-10152
; Sequence 10152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10152

Query Match 34.8%; Score 2403; DB 2; Length 1318;
Best Local Similarity 39.7%; Pred. No. 3.9e-218;
Matches 525; Conservative 238; Mismatches 451; Indels 110; Gaps 26;

QY	2	ABAGLRGWLWLA-----LLRLAQSEPYTHIQPGYCAPYDECG-----KQNELSG 47
DB	32	ASAAALRGHSMTARGALGLLLLLCPAQVFSQ-----SCVWGEGCIAYGDKRYNCEYSG 86
QY	48	SLMTLSNVSCLSNTPARKITGDLHLLLOKICPLRYTGNPTQACCSAKQLVLSLEASISITK 107
DB	87	-----PPKFLPKDGYDLVQLCGLCPFFG-NVSLCCDVRVQLQTLKDNLQLPL 131
QY	108	ALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTR-----VAQIAGAGLPAAVAYEAFYQHS 163
DB	132	QFLSRCPSCFYNNLNFCELTCSPRSQPLNVATEDYDPTVNTQTKTNKELQYYVGS 191
QY	164	FAEQYDSCSRVRVPAATLAVGTTCGVTGYSALCNAQRLNFGQDTGNGGLAPLDIT----- 219
DB	192	FANAMYNACRDVEAPSSNDKALGLLCGKDADA-CNATNWIETMFKNDQCAPFTITPVFS 250
QY	220	---FHLEPGQAVGSGIOPLENGVARNESQGDVATCSQDCAASCAPAIARPQ----- 270
DB	251	DFPVH-----GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPPAW 300
QY	271	---ALDSTFYLGQMPGSLVLIILILCSVFAVVTTILL-----VGFVAPARDKS 314

Qy	1162	SYNAVSLINLVSAVGMSEVFVSHITRGSFAISTKPTWLERAKEATISMGSAVFAGVAMTNL	1221
Db	1189	SLNAVSLVNLVMSGIVSEFCSHITRAFTSTKGRSVRAEEALAHGSSVFSGITLTKF	1248
Qy	1222	PGILVLGLAKAQLIOIPFFRNLITLGLLHGLVFLPVILSYVGGPDVNPAA	1272
Db	1249	GGIVVLAPAKSQIPEIFRMYLAMVLGATHGLIFLPLVLLSYIGPSVYKA	1299
RESULT 4			
US-09-462-136-6			
; Sequence 6, Application US/09462136			
; Patent No. 6426198			
; GENERAL INFORMATION:			
; APPLICANT: Carstee, et al.			
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease			
; FILE REFERENCE: 4239-53894			
; CURRENT APPLICATION NUMBER: US/09/462,136			
; CURRENT FILING DATE: 2000-06-01			
; PRIOR APPLICATION NUMBER: PCT/US98/13862			
; PRIOR FILING DATE: 1998-07-02			
; PRIOR APPLICATION NUMBER: US 60/051,682			
; PRIOR FILING DATE: 1997-07-03			
; NUMBER OF SEQ ID NOS: 13			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 6			
; LENGTH: 1170			
; TYPE: PR1			
; ORGANISM: Saccharomyces cerevisiae			
US-09-462-136-6			
Query Match			
Best Local Similarity 28.8%; Pred. No. 3.4e-116;			
Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;			
Qy	10	LLWAL-----LRLAQSEBYTHIQPGYCAFYDECQNP-----ELSGSLMTLSNVCSLSN	60
Db	3	VLMIIALVGQMLRVQ-----TATCAMYGNCGKSVFGNELPCVPFRSFEPPVLSD	54
Qy	61	TPARKITGDHLLIQLKICPLRYTGPNTQACSAKQVLSLEASITPKALLTRCPACSDNF	120
Db	55	ETSK-----LLVEGGEWEKEVR-YACCTKQDVVALRDLNKAQPLISSCPACKNF	105
Qy	121	VNLCHNTCSNQSLEINVRVAQLGAGQLPVAVYAEFYQHSFASQSDYSCSRVVPAA	180
Db	106	NNLFCHFTCAADQGRFNIKV-EKSKEDKDIVAELDFVFNSSWASEFYDSCKNIKFSAT	164
Qy	181	ATLAVGTMGVYGSALCAQRLNFGQDTGN--GLAPLDTITFHLLEPGQAVSGIQLNE	238
Db	165	N----GYAMDLIGGAKNYSQFLKFLGDAKPMJGGSPFQINYYKIDLANER--KEWQEFND	218
Qy	239	GVARNESQGDVATCSQDCQCAASCPAIAAPQALDSTFYLGOMFG---SLVLIILCSVF	295
Db	219	EVVACDAQ-----YKACSDQSCPHL-KP-LKDGCKVKGPLPCFSLSLVLIIFYTICALF	272
Qy	296	AVVTILVGRVAPARDKSMVDPKGTSLSDK-----LSFSTHT-----LLQPFQCGW	345
Db	273	AFMYVYLCKRKINGAMIVDDDIVPESG-SLDESETNVFESFNNETNPFNGKLANLFTKV	331
Qy	346	TWASWELTILVSVIPVALAAGLV-FTETLTPVELWSAPNSQARSEKAFHDOHGF	404
Db	332	QFSVENPKILITTVSIFVSFIQVATLETPINLWSKSEKKEKEYEDDNFGFP	391
Qy	405	FRNQVILTAPNRSSRYDLSL--LLGPKNFSGIILDLLELLELRLHQLQWSPQA	462
Db	392	YRTQIFVNETGPVLSVETLHWFDVENF-----ITEEL-----QSS	429
Qy	463	RNISLQDTCYAPLNPNTSLVDCINSLLQVFQNNRLLLTANQTLMGQTSQVDWKHF	522
Db	430	ENIGYQDLCPRP-TEDST-----CVIESFTQYFOG-----ALPNKDSWREL	470
Qy	523	LYCANAPLTFKDGITALCSMDYAGVPFPFLAIGYKGDYSEAEALIMTFSLNYPAG	582

Db	471	QECCKFP-----VNCLPTFQQLKTNLL---FSDDDLNAHAFAVVVLTLLTNH---	514
Qy	583	DPRLAQAKWERAFLEEMRAFORRMAGMFOVFTTAERSLEDEINRTAEDLPFATSYIV	642
Db	515	---TQSANWEER-LSEYLLDLKVPBGL-RISFNTEISLEKELNN--NDISTVAISYLM	567
Qy	643	IFLYISIALGSSYSWSRVVDKATIGLGVAVVLGAVMAAMGFFSVGLIRSSILVILQVV	702
Db	568	MFLYATWALRRKDGKTRLL-----LQISGLLIVLASIVCAAGFLTILFGLKSTLIIEVI	621
Qy	703	PFLVLSVGDADNIFIFVLEYQR-LPRRPGEPREVIHIGRALGRVAPSMMLCSLSAICFFLG	761
Db	622	PFLILAIIDINIFLITHEYDRNCEQKPEYSIDQKIISAIGRMSFSLMSLLCOTGCFLIA	681
Qy	762	ALTPMPAVRTFALTSGLAVIDLFLQLMSAFVALLSLSKQEAASRLDVCCKVPQELPPP	821
Db	682	AFVTMPAVHNFAYSTSVIFNGVLQLTAYVLSLSLYEKRSNYKQIT-----	728
Qy	822	GQGBGLLGFQKAYAPFLHMTIRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDS	881
Db	729	-GNEETKESFLKTYFKMLTQ---KRLIIIFSAWFFTSILVELPEIQFGLDQTLAVPQDS	784
Qy	882	YLLDYFLNARYPEVGAPVYFVTTLGVNFSSEAGMNAICSS-AGCWNFSFTQKIYATEP	940
Db	785	YLVDFYKDVYSFLNVGPPVYMWVK-NLDLTRQNOQKICGKFTTCERDSLAVLIE---QE	840
Qy	941	PEQSYLAIPASSWYDDFDIMLTSS--CCRLYISGPNKDKFCPSTVNSLNCNKMCSITM	998
Db	841	RHRSTITEPLANLWLDYFMFLNPQDQCCRL---KKGTDVCPSPFSRRC-ETCFQ--Q	894
Qy	999	GSVR-----PSVQBFHKYLPFWFLNDRPNIKCPKXGLAAYSTSVNLSDGOVLASRFMAY	1052
Db	895	GSWYNWMSGFPEGKDFMEYLSIWIN-APSDPCPLGGRAPYSTAL-VYNETSVSASFRTA	952
Qy	1053	HKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITVNFYQYLITLPEG	1112
Db	953	HHPLRSQKDFQAV-----SDGVRISSPFPELDMFAYSPPFYFFVQYOTLGPJT	1001
Qy	1113	LFMLSLCLVPTFAYSCLLGLDLRSLNLSIVMILVDTVGFMALMDISYNVSLNLIV	1172
Db	1002	LKLGSAIILFFISSVFL-QNIRSEFLALVVTMIIIVDIGMALLGISLNAVSLVNI	1060
Qy	1173	SAVGMSEVFVSHITRSPAI---STKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLQ	1229
Db	1061	ICVGLGVFEVCHIVESFTVVPSETKDNANRSLVSLNTIGESVIKGITLTTFIGVCVLAF	1120
Qy	1230	AKAQIQLIFFFRNLITLGLLHGLVFLPVILSYVG	1266
Db	1121	AQSKI FDFYFRMFWFTLIIVAALHALLFLPALLSLFG	1157
RESULT 5			
US-09-462-136-9			
; Sequence 9, Application US/09462136			
; Patent No. 6426198			
; GENERAL INFORMATION:			
; APPLICANT: Carstee, et al.			
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease			
; FILE REFERENCE: 4239-53894			
; CURRENT APPLICATION NUMBER: US/09/462,136			
; CURRENT FILING DATE: 2000-06-01			
; PRIOR APPLICATION NUMBER: PCT/US98/13862			
; PRIOR FILING DATE: 1998-07-02			
; PRIOR APPLICATION NUMBER: US 60/051,682			
; PRIOR FILING DATE: 1997-07-03			
; NUMBER OF SEQ ID NOS: 13			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 9			
; LENGTH: 1296			
; TYPE: PR1			
; ORGANISM: Caenorhabditis elegans			
US-09-462-136-9			

Db 24 PGRAGGRRRTTGGRRRAAADPRD-----YLHRPSYCDAAFALEQI--- 65
Qy 285 LVLIILILCSFVAVVTILLVGRVAPARDKSRKMDPK-----RGTSLSDKLSPSTHTLL 337
Db 66 -----SKGKATGKAPLWLRKQFRLFKLGCYIQKNC 98
Qy 338 GQFFOGWGTWASWPLTILVSVIPVVALAGLFTTDPVLSNPNQARSEKAFH 397
Db 99 GKF-----LVGLLIFGAFVGLKAAANLETNVEELWVEVGGVRSRELYT 143
Qy 398 DQHGFPPFTN-QVILTAPNRSSRYDSILLGPKNFSGILDLLELE--LQERLHL 454
Db 144 ROKIGEEAMFNQLMIQPKERG-----ANVLTEALLQHLDSALQASRVH 190
Qy 455 QWSPQEAQRNTSLQDICAP-----LNPNTSLYDCCINSLLQYFQNNRTLLTA 505
Db 191 YMYN-----ROWKLEHLYKSGBELITETGMDQ1IEVLYPCLIIITPLDCFWGAKLQSGTA 246
Qy 506 NOTLMGO-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFDPLEFLEBELKINQVDSWEEMLNKAIEVGHYMDRPLCPADPDC 304
Qy 528 -----APL-----TFKGT-----ALALSQWA 544
Db 305 PATAFNKSTKPLDMALVNLGCHLSRKYMWQBELIVGGTVKNSGKLSAHALQTMF 364
Qy 545 DYCAPVFPFLAIGYKGYDBAEALIMTFSLNYPAGDPRLQAQAKLWEEAFLEMRAPQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINNE---DKAAAILLEAQRTVEVHVQSV 411
Qy 605 RMAGHFQVTFARSLDEINRTAEDLPIFATSYIVIFLISALGSYSWSRVMVDS 664
Db 412 AQNSTQKVLSTFT--TTLDILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
Qy 665 KATLGGVAVVLGAYMAAGPFSYLGIRSLVLOVPFLVLSVGNADNIFVLEYORL 724
Db 466 QGAVGLAGVLLVALSVAGLGLCSLIGISFNATTOVLFFLALGVGVDDVFLAHAFST 525
Qy 725 PRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLIGALTMPAVRTFALTSGLVILDF 784
Db 526 GQNKRIPTEDRTGECRLKRTGASVALTSIGNVTAFFMAALIPALRAFSLQAAVVVFNF 585
Qy 785 LLOMSAFVALLSDSKROASRLDVCC-----VKQEL-----PPG 822
Db 586 AMVLLIFPAILSMDLRYRDRDLDFCCFTSPCVSRVIVQEPQAYTDTHTNRYSPPPY 545
Qy 823 QGEG----- 827
Db 646 SSHSFAHETOITMQSTVOLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGPFQKAYAPFLHWTIRGVVLLFLALFGVSL 861
Db 706 SSTRDLLSQFSSSLHCLPEPCTKWTLSSPAETHYAPFLKPKAKVWVVFPLFGLLGVSL 765
Qy 862 YSMCHISVGLDQELAPKDSYLDYFLFNRYEYVGVAPYVFTVTLGNFSSSEAGMAICS 921
Db 766 YGTRVRDGLDLDIVPRETREYDFIAQFKYFSF-----YNM----- 803
Qy 922 SAGCNFSFTQKQYAT-----EFPEQSYLAIPASS-----WVDDFIDL----- 961
Db 804 -----YIVTQKADYFNIQHLLYDLHRSFSNVKYVWLEENKQLPKMHLHYFRDMLQGLQD 857
Qy 962 -----TPSS-----CRLVIGSNKDKFCPSTVNSLNCNCSMTGMS 1000
Db 858 AFDSDETQKIMPNYKNGSDGVLAYKLLVQTGRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFKYL--PWFLLND-----RPNIKCPKGLAAY--STSVNLTSQDG 1043
Db 914 INPSA--FYIYLTAVNSNDPVYAASQANTRPHRPEWHDK--ADTMPETRLRIPAE 968
Qy 1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPGTDPAPFVPPYITNVFY 1102
Db 969 IEYAQFPFLNGLRDTSDFVEAIEKVRKITSNYTSLGLSSYPNG-----YPF-----LFW 1018

Qy 1103 EOYLITLPEGLFMLSCLVPTRAVSCLLGLDLRSGLNLLSIWMLIVDTVGFMAWDIS 1162
Db 1019 EQYIGRHWLULFISVVVLAFTFLVCAVFLNPNWTAGII-VWVLAATVVELFGWGLIGIK 1077
Qy 1163 YNAVSLINLVSAGMGSVEFVSHITRSF--AISTKPTWLERAKATISMGSAGVAGVAMTN 1220
Db 1078 LSAVPPVILLIASGVIGVEFTVHVALAFLTAIGDKN---RRAVLALEHMFAPVLDG-AVST 1133
Qy 1221 LPGLILVGLAKAQIQLIFFRNLNLLITLLGLLHGLVPLPVLISYVG--PDVNPALALEQ- 1277
Db 1134 LLGLVMLAGSEFDIVRYFVAVLAILTILGLVGLNGLVLLPVLSSFFGYPEVSPANGNLRL 1193
Qy 1278 --KBAEAAVAVMVASCPNHPKSRVSTADNITYVNHSPFGSING 1317
Db 1194 PTPSPPEPPSVVRFPMPGHTH--SGSDSDSESVSQTTVSG 1233

RESULT 7
US-08-656-055-19
; Sequence 19, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
Qy 225 PGQAVSGIQPLNEGVARNCSQGDVATCTCQDCAASCPAIPARQALLSTFVLGMPGS 284
Db 24 PGRPAGGRRRTTGGRRRAAADPRD-----YLHRPSYCDAAFALEQI--- 65
Qy 285 LVLIILILCSFVAVVTILLVGRVAPARDKSRKMDPK-----RGTSLSDKLSPSTHTLL 337

Db 66 -----SKGKATGRKAPLWLRKAFORLLFKLGCYIQKNC 98
Qy 338 GQPFQGWGTWASWPLTILVSLVPVVALAAGLVFTELTDPVELWSAPNSQARSEKAPH 397
Db 99 GKF-----LVVGLLIFGAPAVGLKAAANLETVBELVVEVGGVRSRELYT 143
Qy 398 DQHFQFPFRFTN-QVILTAPNRSSRYRDSLLGPKNFGSGLDLDLLELELE--LQERLRLH 454
Db 144 RQKIGSEAMFNPMQITQPKBEG-----ANVLTTBALOHLDSALQASRVHV 190
Qy 455 QVMSPEAQRNISLQDIQYAP-----LNPDTSLYDCINCISLLQYFQNNRTLLLLTA 505
Db 191 YMN-----RQMKLEHLCKYSGELITETGYMDQIIEVLYPCLITPLDCTWEGAKLQSGTA 246
Qy 506 NOTLMQO-----TSQVDMKDHPLY-----CAN-----527
Db 247 --YLLGKPLRWTFNFDPLFLEBELKKINYQVDSWBEMLNKAEGVGHGMDRPCLPNADPDC 304
Qy 528 -----APL-----TFKDOT-----ALALSQWA 544
Db 305 PATAPNKNSTKPLDMALVNLGCGHLSRKYMHQWBEELIVGGTVKNSGTGLVSAHALQTMF 364
Qy 545 DYCAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAOKLWEEAFLEEMRAFO 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINNE-----DKAAALEAMORTTVEVVHQSV 411
Qy 605 RRMAGMFQVPTAERSLEDEINRTAEDLPIFATSYIVIFLYXISLALGSYSMSRWVDS 664
Db 412 AQNSTQKVSFTT--TTLDILKSPSDSVIRVASGYLLMLAVACLTM---LRWD--CSKS 465
Qy 665 KATLGLGGVAVVLGVAWMAANGFSYLGIRSSLVLOVPFLVLSUGADNIFPVLEYQRL 724
Db 466 OGAVGLAGVLLVALSVAAGLGCSLIGISFNAATTQVLPFLALGVGDDVFLLAHAFSET 525
Qy 725 PRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTTALTSGLAVILDP 784
Db 526 GQNKRIFFEDRTGECLEKRTGASVALTSISNTAFFMAALIPALKAFAQLQAAVVVNF 585
Qy 785 LQMSAFVALLSLSKROBSRLDVCCC-----VKPQEL-----PPPG 822
Db 586 AMVLLIFPALLSMDLYRREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHTNRYSPPPY 645
Qy 823 QOEGL-----827
Db 646 SSHSFAHETQITWQSTVQLRTBYDPHTHYVYTAEPRISEISVQPVTVTDTLSCQSPEST 705
Qy 828 -----LLGFFOKAVAPLHWTIRGCVLLLFALFGVSL 861
Db 706 SSTRDILLSQFSSSLHCLPEPCTKWTLSFAEKHYAPFLKPKAKVWVIFLGLLGVSL 765
Qy 862 YSMCHI SVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTILGYNPSSBAGNNAICS 921
Db 766 YGTVTRVDRGLDITDIVPRETRYDFIAQPKFSP-----YNN-----803
Qy 922 SAGCNFFSTQKIQVAT-----EPPEGSYLAIPASS-----WVDDFDIWL-----961
Db 804 -----YIVTQADYPNIGHLLYDLHRSFSPNVKYVMLEENKQLPKMWLHYFRDMLQGLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKPCSTVNSLNCNKMCMITWGS 1000
Db 858 AFDSWETGKIMPNNKNGSDGVLAYKLLVOTGSRDK--PIDISOLT--KORLVADAGI 913
Qy 1001 VRPSVEQFHXYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYIYLTAWSNDPVAYASQANIRPHREWHDK---ADYMPETRLRIPAEP 968
Qy 1044 VLASRFMAHYHKLKNSQDYTEALRAARELAANITA--DLRKVPGTDPAPFVPPYTTINVPY 1102
Db 969 IEYAQPPFPLNGLRDTSDFVEAIEKVRTICSNYTSLSGLSSYPNG-----YPF-----LFW 1018
Qy 1103 EOYLTLPEGLPMLSLCVPTFAVSCLLGLDLRSLGNLLSIVMLVDTVCGFMAWDIS 1162
Db 1019 EOYIGLRHWLLFISVVLACTFLVCAVFLNPNWTAGII--VMVLAJMTVELFGMGLIGIK 1077

Qy 1163 YNAVSLINLVSAVGMSEVFEVSHITRSF--AISTKPTWLBRAKEATISMGSVAVFAGVAMTN 1220
Db 1078 LSAVPVILLIASVGIGVEFTVHVALAFLTAIGDKN---RRAVLALHEMFAPVLGD-AVST 1133
Qy 1221 LFGILVLGAKAQLOIQIFPFRNLNLLITLGLLHGLVPLVILSYVG--PDVNPAALAEQ- 1277
Db 1134 LLGLVMLAGSBDFIVRYFFAVLAILITILGVLNGLVLLPVLLSFFGYPYDEVSPANGLNRL 1193
Qy 1278 --KRAEEAAVAAVMVASCPNHPRSVSTADNIYVNHSPESGSIK 1317
Db 1194 PTPSEPPEPPSVVRPAMPPOGTH--SGSDSDSEYSSQTTVSG 1233

RESULT 8

US-08-954-668-19
; Sequence 19, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,668
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-668-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PQAVGSGIQPLNEGVARNCSQGGDVATCSQDCAASCPAIPARQALDSTFYLGQMPGS 284
Db 24 FGRPAGGRRRTTGLRRAAADPRD-----YLRHRSYCDAAAFALQI--- 65
Qy 285 LVLIILICSVAVVTILLVGRFVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAFORLLFKLGCYIQKNC 98
Qy 338 GQPFQGWGTWASWPLTILVSLVPVVALAAGLVFTELTDPVELWSAPNSQARSEKAPH 397
Db 99 GKF-----LVVGLLIFGAPAVGLKAAANLETVBELVVEVGGVRSRELYT 143
Qy 398 DQHFQFPFRFTN-QVILTAPNRSSRYRDSLLGPKNFGSGLDLDLLELELE--LQERLRLH 454

Db	144	RQKIGEEAMFNPQMIQTPEKEG-----ANVLTEALLQHLDSALQASRVHV	190
Qy	455	QVWSPQARNISLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA	505
Db	191	YMYN-----RQWLEHLCYSGELITETGYMDQIIEYLYPCLIIITPLDCFWEAKLQSGTA	246
Qy	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----	527
Db	247	--YLLGKPELRWTFDPLELEELCKINYQVDSWEEMLNKAEVGHGYMDRPLNADPDC	304
Qy	528	-----APL-----TFKOGT-----ALALSCMA	544
Db	305	PATAPNKNSTKPLDMALVINGCGHLSRKRYMHWOBELIVGGTVKNSTGKLVSAAHALQTMF	364
Qy	545	DYGAPVFPFLAIGGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKWEEAFLEEMRAFQ	604
Db	365	QLMTPKQMY---EHPKGYEY---VSHINWNE---DKAAAILBAMQRTYVEVHHQSV	411
Qy	605	RMAGMFQVTFPAAERSLEDEINRTTAEDLPFIPATSYIVIFLYISALGSYSSWSRVWVDS	664
Db	412	AQNSTQKVLSPFT--TTLDILKSPFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS	465
Qy	665	KATLGLGGVAVVLGVAMAAMGFPSYLGIRSSLVILQVPPFLVLSVGAADNIFIFVLEYQRL	724
Db	466	QGVAGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAFSET	525
Qy	785	LLQMSAFVALLSLDKRQASRLDVCCC-----VKPQEL-----PPPG	822
Db	586	ANVLLIFPALLSMDLYRREDRLDFCCFTSPCVSRVIOVBPQAYTDTHDNTRYSPPPPY	645
Qy	823	QGBGL-----	827
Db	646	SSHSPAHETQITWQSTVQLRTEYDPHTHVYTTAEPRSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFFQKAYAPFLLHMTIRGVVLLFLALFGVSL	861
Db	706	SSTRDILLSQFSSSLHCLBPPCTKTWLTSSFAEKHYAPFLKPKAKVWVIFLGLLGVSL	765
Qy	862	YSMCHISVGLDQBLALPKQSYLLOYFLFNRYEFGAPVYFVTTLGNFSSSEAGMAICS	921
Db	766	YGTTRVRDGLDITDIVPRETREYDFIAAQKYFSF-----YNM-----	803
Qy	922	SAGNNFSPQTKQYAT-----EPPEQSYLAIPASS-----WVDDFDIDLW----	961
Db	804	-----YIVTQKADYPNIQHLLYDLHRSFNSVNYVMLEENKQLPKMWLHYFRDWLQLOQD	857
Qy	962	-----TPSS-----CCRLYISGNKDKFCPSTVNSLNCNKCWSITMGS	1000
Db	858	AFDSDWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVDADGI	913
Qy	1001	VRPSEVQFHXYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLTSQDG	1043
Db	914	INPSA--FYIYLTAWSNDFVAVAAISOANIRPHRPWWHDK---ADYMPETRLRIPAEEP	968
Qy	1044	VLASRFWAYHKPLKNSODYTEALRAARELAANITA--DLRKVPGTDPAPFEPFVYITNVFY	1102
Db	969	IEYAAQFPFYNLGLRDTSDFVAIEKVRTICSNYTSLSGSSYPNG-----YPF-----LFW	1018
Qy	1103	EQYLTILPEGLFMLSCLVPTFAVSCLLLGLDLRSLGNLLSIVMILVDTVGFMALWDIS	1162
Db	1019	EQYGLRHWWLLFISVVLACTFLVCAPFLNWPATAGII--VMVALMTVELFGMMGLIGIK	1077
Qy	1163	YNVSLINLNVSAVGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1078	LSAVPWWIIASVSGIVGEFTVHVVALAFLTAIGDKN---RRAVTALEHMFAPVLDG-AVST	1133
Qy	1221	LPGLIIVLGLAKAQILQIFPFRNLNLLITLLGLLHGLVFLPVILSVYG--PDVNPALALEQ-	1277

Db	1134	LLGVLMAGSEPDFIVRYFFAVVLAAILTILGLVINGLVLPLVLLSFFGFPYPEVSPANGLNRL	1193
Qy	1278	--KRAEEAANAAMVASCNPHPSRVSTADNIYVNVHSPGSGIKG	1317
Db	1194	PTSPSPPPSVVRFAMPGGHTH--SGSDSDSDSYSSQTTVSG	1233
RESULT 9			
US-09-268-140-5			
; Sequence 5, Application US/09268140			
; Patent NO. 6268176			
; GENERAL INFORMATION:			
; APPLICANT: Gemmill, Robert M.			
; APPLICANT: Drabkin, Harry A.			
; TITLE OF INVENTION: TRC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR, PATCHED			
; FILE REFERENCE: 93445-00004			
; CURRENT APPLICATION NUMBER: US/09/268,140			
; CURRENT FILING DATE: 2000-03-12			
; PRIOR APPLICATION NUMBER: US 60/077,723			
; PRIOR FILING DATE: 1998-03-12			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: Patentin Ver. 2.0			
; SEQ ID NO 5			
; LENGTH: 1447			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-268-140-5			
Query Match 9.4%; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVSGIGIPLNEGVARCNESQGDVATCSCQDCAASCPAIAFPQALDSTFYLGQMPGS	284
Db	24	PCRPAAGGRRRTTGLRRAAAPDRD-----YLHRFSYCDAAFALEQI---	65
Qy	285	LVLIILGCVFAVVTILLVGRFVAPARDKSKWDPK-----KGTSLSDKLSFSTHTLL	337
Db	66	-----SKKATGRKAPLWLRKQRLFLKLCGYIQKNC	98
Qy	338	GQFFGCGWTWASWPLTLLVLSVIPVVALAGLVFTELTDPVELWSAPNSQARSEKAPH	397
Db	99	GKF-----LVGLLIFGAPAVGLKKAANLETNVEELWVEVGGVRSRLNVT	143
Qy	398	DOHFQFPFRTN-QVILTAPNRSSRYRDSILLGPKNFSGILDDLDLLELE--LQERLRLH	454
Db	144	RKIGEEAMFNPQMIQTPEKEG-----ANVLTEALLQHLDSALQASRVHV	190
Qy	455	QVWSPQARNISLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA	505
Db	191	YMYN-----RQWLEHLCYSGELITETGYMDQIIEYLYPCLIIITPLDCFWEAKLQSGTA	246
Qy	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----	527
Db	247	--YLLGKPELRWTFDPLELEELCKINYQVDSWEEMLNKAEVGHGYMDRPLNADPDC	304
Qy	528	-----APL-----TFKOGT-----ALALSCMA	544
Db	305	PATAPNKNSTKPLDMALVINGCGHLSRKRYMHWOBELIVGGTVKNSTGKLVSAAHALQTMF	364
Qy	545	DYGAPVFPFLAIGGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKWEEAFLEEMRAFQ	604
Db	365	QLMTPKQMY---EHPKGYEY---VSHINWNE---DKAAAILBAMQRTYVEVHHQSV	411
Qy	605	RMAGMFQVTFPAAERSLEDEINRTTAEDLPFIPATSYIVIFLYISALGSYSSWSRVWVDS	664
Db	412	AQNSTQKVLSPFT--TTLDILKSPFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKS	465
Qy	665	KATLGLGGVAVVLGVAMAAMGFPSYLGIRSSLVILQVPPFLVLSVGAADNIFIFVLEYQRL	724
Db	466	QGVAGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAFSET	525
Qy	725	PRRGPPEVHIGRALGRVAPSMLLCSLSEAI-CFFLGLALTTPMPAVRTFALTSGLAVIDLF	784

Db 526 GQNKRLPFEDRTGECUKRTGASVALTISNVTAFMAALIPALRAFSLQAAVVVVFNF 585
QY 785 LLOMSAFVALLSDSKRQBSRLDVCC-----VKQEL-----PPPG 822
Db 586 AMVLLIFPAILMDLYRDRRLDIFCCFTSPCVSRVQVPEQAYTDHNTRYSPPPY 645
QY 823 QSEGL----- 827
Db 646 SSHSFAHQITQWSTVQLRTEYDPRHVVYTTAEPREISVQPVTVTQTLSCQSPST 705
QY 828 -----LLGFQKAVAPFLLHWTGRCVLLLLFLALFGVSL 861
Db 706 SSTRDLLSQFSSSLHCLPCCPKWTLSFAEKHYAPFLPKKAVVWVIFLGLLGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLNFRPEVGAPVYFVTTLGFNFSSEAGNNAICS 921
Db 766 YGTRVRDGLDITDIPRETREYDFTAAQFKYF-----YNN----- 803
QY 922 SAGCNPFSTQKIYAT-----EPESYLAIPASS-----WVDDFDWL----- 961
Db 804 -----YIVTORADYENIQHLLYDLHRSFSNVKYVMLEENKQLPKMWLHYFRDMLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKCPSVTNSLNCCLKMCSITMGS 1000
Db 858 AFDSWETGKIMPNYKNGSDGVLAYKLLVOTGSRDK--PIDISQLT--KQRLVDADGI 913
QY 1001 VRPSVEQHKYL--PWFLND-----RPNKCPKGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYILTAVNSNDPVAYASQANIRPHREWHDK--ADYMPETRIRIPAAEP 968
QY 1044 VLASRPMATHKPLKNSQDYTEALRAARELAANITA--DLAKVPGTDPAPFVFFYTTINVFY 1102
Db 969 IEYAQPFYINGLRDTSDFEAEIKVRTICSNVTSGLSSYPNG-----YPF-----LFW 1018
QY 1103 EOYLTLPEGLFWLSICLPTFAVSCLLGLDRLSLNLLSIVMLLVDTVCFMALWDIS 1162
Db 1019 EYQIGLRHLLFISVLACTFLVCAVFLNFWTAGII--VMVALMTVELFCMGGLIGIK 1077
QY 1163 YNAVSLINLVSAGMSVEFVSHITSF--AISTKPTMLERAKEATISMGSAVFAVAMTN 1220
Db 1078 LSAPVVLIIASGVGEVETVHVALAFLTAIGDKY--RRVLALEHMPAPVLDG-AVST 1133
QY 1221 LGLVLGLAKAQLQIIPFRLMLLITLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
Db 1134 LGLVLMLAGSEPDFIVRYFPAVLAILTLGLVNLGLVLLPVLISFGFPYEVSPANGNLRL 1193
QY 1278 --KRAEEAVALVAVASCPNHPRSVSTADNIYNHSPGSIK 1317
Db 1194 PTPSPPPSVVRFPMPGPHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 10

US-08-918-658-19

; Sequence 19, Application US/08918658

; Patent No. 6429354

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/918,658

; FILING DATE: 22-Aug-1997

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/656,055

; FILING DATE: <Unknown>

; APPLICATION NUMBER: 08/540,406

; FILING DATE: <Unknown>

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1447 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-08-918-658-19

Query Match

Best Local Similarity 9.4%; Score 651.5; DB 2; Length 1447;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQQAVSGIQPLNEGVARNCSQGDVATCSCQDCAASCPAIPARQALDSTFVLGQMPGS 284
Db 24 PGRPAGGRRRTTGLRAAAPDRD-----YLRPSYCDAAFALEQI--- 65
QY 285 LVLIILCSVFAVVTILLVGFVAPADKSKRWDPK-----KGTSLSDKLSFSFHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAFQRLFLKLCYIQKNC 98
QY 338 GQFFQGWGTWASWPLTILVLSVIPVVALAAGLVTELTTPVELWSAPNSQARSEKAFH 397
Db 99 KGL-----LVVGLLIFGAPAVGLKKAANLENTNVEELNVEVGGVRSRELNT 143
QY 398 DQHGFFPFTN-QVILTAPNRSSVRYDSLLGPKNFSGILDLDLLELE--LOERLRLH 454
Db 144 RQKIGEEAFNPQLMIQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QVMSPEAQRNLSQDICYAP-----LNPDNLSYDCCINSLLQVFNRTLLLLTA 505
Db 191 YMYN----RQWKLHLCYKSGBELITETGYMDQIIEYLYPCLITPLDCFWEGAKLQSGTA 246
QY 506 NQTLNGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 ---YLLGKPLPRTNPDPLELEELKKINYQVDSWEEMLNKAEVGHYMDRPLNFPADPC 304
QY 528 -----APL-----TFKQGT-----ALALSCWA 544
Db 305 PATAPKNSTFPLDMALVINGCHGLSRKYMHWQBELIVGGTVKNSTKGLVSAHALQTFW 364
QY 545 DYGAVPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEBMRAPQ 604
Db 365 QLMTFPKQWY---EHFKGYEY-----VSHINWNE---DKAAAILAEMQRTVEVVHQS 411
QY 605 RRMAGMFQVTTAERSLEDEINRTAEDLPITPATSYIVIFLYISALAGSYSSWSRWVDS 664
Db 412 AQNSTQKVLSTPT-TTLDDILKSFSDSVIRVASGYLLMLAYACITM---LRWD--CSKS 465
QY 665 KATLGLGVAVVVLGAVMAAMGFPSVLGRSSILVILQVVPFLVLSVGADNIFLVEYQRL 724
Db 466 QGAVGLAGVILVALSVAAGLGLCSLIGISFNAAATQVLFPFALGVGVDDVFLAAHAFSE 525
QY 725 PRPGEPREHVHGRALGRVAPSMLLCSLSEAIKCFPLGALTTPMPAVRTFALTSLGLVILDF 784

526 GQNKRIFFEDRTGELKRTGASVALTSISNVTAFPMALIPALRAFSLQAAVVVVVFN 585
785 LLQMSAFVNLISLDSKQESRLDVCC-----VKPQEL-----PPPG 822
586 AMVLLIFPAILSMOLYRRDRRLDFCCFTSPCVSRVQVEPQAYTDTHTNTRYSPPPY 645
823 QGEG-----827
646 SSHSFAHETQITMQSTVQLRTEYDPHTHVYTTAEPRSEISQVPVTVTQDTLSQCSPEST 705
828 -----LLGPKQAYAPPELLHWITRGVWVLLLFALFCVSL 861
706 SSTROLLSQFSDSLHCLPEPCTKTWLTSSFAEKYAPFLPKAKVWVVFILGILGVS 765
862 YSMCHISVGLQELALPKDSVLLDYFLPLNRYFVEGAPVYFVTTILGYNFSEAGNAICS 921
766 YGTRVRDGLDITVPRETREYDFIAQPKYF-----YNN-----803
922 SAGCNFSFTQKIQYAT-----EPFQSYLAIPASS-----WVDDFIDWL-----961
804 -----YIVTKADYPNIQHLLYDLHRSFNKVKYVMLEENKQLPKWMLHYFRDMLQGLQD 857
962 -----TPSS-----CCRLYISGPNKDKFCPSVNSLCLNCKNCSITMGS 1000
858 AFDSDWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
1001 VRPSVEQPHKYL--PWFLND-----RPNIKCPKGGLAAY--STSVNLITSDGQ 1043
914 INPSA--FYIYTAVUSDNPVAYASQANIRHREWHDK--ADYMPETRLRIPAAEP 968
1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPGTDPAEVPPYITNVFY 1102
969 IEYAQPPFYGLNGLRDTSDFEAEIKVRTICSNYSGLSSYPNG-----YPF-----LFW 1018
1103 EOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSLNLLSIVMLVDTVGFMALWDIS 1162
1019 EQVIGLRHLLLFISVLAFTFLVCAVFLNLPWTAGII--VMVLALTEHMFAPVLDG-AVST 1133
1163 YNAVSLINLSVAGMSVFPVSHITSF--AISTKPTWLERAKEATISMGSVAFAGVAMTN 1220
1078 LSAVPVILISVIGVGEVTVHVALAFLTAIGDKN--RRAVLALHMFAPVLDG-AVST 1133
1221 LPEGLVGLAKAQIQLIFFRNLNLTLLGLHLGLVFLPVILSYVG--PDVNPALALEQ- 1277
1134 LLGVLMAGSEDFIVRYEFAVLAILTILGVLNGLVLLPVLISFFGYPYEVSPANGLARL 1193
1278 --KRAEAVNAVVASCPNHPHSRVSTADNIYVNHSPESIGK 1317
1194 PTPSPPPPPVVRFAFMPPGHHT--SGSDSDSEISSQTTVSG 1233

RESULT 11
US-09-724-631-19
Sequence 19, Application US/09724631
Patent No. 6551782
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
JOHNSON, RONALD L
JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
CORRESPONDENCE ADDRESS: 19
ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/724,631
FILING DATE: 28-No. 6551782-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/656,055
FILING DATE: 1996-05-31
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
STRANDEDNESS: single
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-724-631-19
Query Match 9.48; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PGQAVSGIOPLEBVGARCNESQGDVATCSQCAASCPAARPOLADSTFYLGQMPGS 284
DB 24 PGRPAGGRRRTTGLRRAAPDRD-----YLHRPSYCDAAFALEQI--- 65
QY 285 LVLIILICSVFAVVITILLVGRVAPARDKSRMDPK-----KGTSLSDKLSFSTHTLL 337
DB 66 -----SKGATGRKAPLMRAKFORLLFKLGCYIQKNC 98
QY 338 QQFQGGTGWASWPLTILVLSVIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFH 397
DB 99 GKF-----LVVGLLIFGAFVGLKAANLETVEELWVEVGRVSRRLNYT 143
QY 398 QHFGPPFRTN-QVILTAPNRSSRYDLSLLGPKNFSGILDLDLLELLE--LQERLRLH 454
DB 144 RQKIGEEAMFNPOLMIQTPKEG-----ANVLATTEALLOHLSALQASRVHV 190
QY 455 QWMSPEAQRNLSQDICIYAP-----LNPNTSLYDCINSILQYFQNNRTLLTLTA 505
DB 191 YMYN-----RQWKLHLCKYSGELITETGYMDQIIEYLYPCLIIITPLDCFWEKALQSGTA 246
QY 506 NOTLMGQ-----TSQVMDKDHFLY-----CAN-----527
DB 247 --YLLGKPLPLWNTNPDPLFLEELKKINYQVDSWEMLNKAEBVGHGYMDRCLNPADPDC 304
QY 528 -----APL-----TFKDG-----ALALSQWA 544
DB 305 PATAFNKSTKPLDMALVNLGCHSLRKMYHQBELIVGGTVKNSGKLSAHALQTMF 364
QY 545 DYGAVPFPPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFQ 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE-----DKAAAILLEAWQRTTVEVHQSV 411
QY 605 RMAGNFQVFTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISALGSSYSWSRVMDVS 664
DB 412 AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMLAACLTM---LRWD--CSKS 465
QY 665 KATLGLGVAVVILGAVMAAMGFFSYLGRSSIVILQVFPFLVLSVGCADNIFIFVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGVDDVFLAHAFSET 525
QY 725 PRPGEPRVHIGRALGRVAPSMLLCSLSEALCFGLGALTMPAVRTFALTSLGVLIDF 784
DB 526 GQNKRIFFEDRTGELKRTGASVALTSISNVTAFPMALIPALRAFSLQAAVVVVVFN 585

Db	646	SSHFAHETQITMQSTVQLRTEYDPHTHYTTAEPRSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFFQKAYAPFLLHWITRGVVLFLFALFGVSL	861
Db	706	SSTRDLLSQFSDSLHCLPEPCTKWTLSFAEKHYAPFLKPKAKVWVIFLFLGLGVSL	765
Qy	862	YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFTTLGYNFSSEAGHNAICS	921
Db	766	YGTTRVRDGLDLDIVPRETREYDFIAAQKYFSF-----YNN-----	803
Qy	922	SAGCNPFSTQKIQYAT-----EPQSYLAIPASS-----WDDFDIDL-----	961
Db	804	-----YIVTQADYPNIQHLLDYLRHSFNKVKYVLEENKQPKMWLHYFRDMLQGLQD	857
Qy	962	-----TPSS-----CCRLYISGPNKDKFCPSTVNSLCLNCKMCSITMGS	1000
Db	858	AFDSDWETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVDADGI	913
Qy	1001	VRPSVEQFHYL--PMFLND-----RPNKCPKGLAAY--STSNNLTSDQG	1043
Db	914	INPSA--FYIYLTAWNSNDPVAYAASQANIRPHREPWVHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASRFMAHYKPLKNSQDYTEALRAARELAANITA--DLRKVPQTDPAFEVFPYTTINVFY	1102
Db	969	IYEAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSGLSSYPNG-----YPP-----LFW	1018
Qy	1103	EQYILRLPGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVVGFMALWDIS	1162
Db	1019	EQYIGLRHMLLFISVVLACTFLVCVAFLLNPWTAGII--VMVLALMTVELFGMGLIGIK	1077
Qy	1163	YNAVSLINLVSAGVSVFVSHITRSF--AISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1078	LSAAPPVILLIASVGIGVEFTVHVALAFLTAIGDN---RRVLALEHMFAPVLDG-AVST	1133
Qy	1221	LPGLVLGLAKAQIQLIFFRMLNLITLGLHGLVFLFVILSVYG--PDVNPALALEQ-	1277
Db	1134	LGVLMLAGEPDFIVRYFAVLAIIITLGLVNLGLVLLPVLSFFGYPYEVSPANGLNRL	1193
Qy	1278	--KRAEEAANAAMVASCNPHPSRVSTADNIYNHSPGSIKG	1317
Db	1194	PTSPSPPPSVVRFAAMPQGH--SGSDSDSEYSSQTTVSG	1233
RESULT 13			
US-09-807-007-6			
; Sequence 6, Application US/09807007			
; Patent No. 6881833			
; GENERAL INFORMATION:			
; APPLICANT: ZAPHIROPOULOS, Peter et al.			
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY			
; FILE REFERENCE: 2921-0130P			
; CURRENT APPLICATION NUMBER: US/09/807,007			
; CURRENT FILING DATE: 2001-04-06			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: Patent in version 3.1			
; SEQ ID NO 6			
; LENGTH: 1447			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-807-007-6			
Query Match 9.48; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.08; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVGSGIOPLNGBVARNCSQDDVATCSODCAASCFAIARPOALDSTFYLGQMPGS	284
Db	24	PGRPAGGRRRTTGLRAAAPRED-----YLHRPSYCDAAFALEQI---	65
Qy	285	LVLIIILCSFVAVVITLLVGRVAPARDKGMVDPK-----KGTSLSKLSFSTHLL	337
Db	66	-----SKGATGRKAPLMRAKFORLLFKLGCYIQKNC	98

Db 1078 LSAPVWIIASVGVGVEFTVHVALAFLTAIGDN---RRVLALEHMFAPVLDG-AVST 1133
Qy 1221 LPGLVLGLAQLQIQIFFRMLNLITLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
Db 1134 LLGVLMLAGSEDFVIRYFFAVLAIIITLGLVNLGLVLLPVLSPFGPYEVSFANGNLRL 1193
Qy 1278 --KRAEAVAAMVWASCPNHPRSRVSTADNIYNHSPGSIKG 1317
Db 1194 PTPSPPPSVVRFAMPFGHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 14

US-09-754-032-19

; Sequence 19, Application US/09754032

; Patent No. 6921646

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; ; GOODRICH, LISA V

; ; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/754,032

; FILING DATE: 03-Jan-2001

; CLASSIFICATION DATA:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/08/540,406

; FILING DATE: 06-Oct-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 19:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1447 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 19:

US-09-754-032-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

Qy 225 PQQAVSGIQLNEGVARNESQGDVATCSQDCAACPALARPQALDSTFYLGQMPGS 284

Db 24 PGRPAGGRRRTGGLRAAADPRD-----YLHRSYCDAAFALEQI--- 65

Qy 285 LVLIILILCSFAVAVTILLVGRFVAPARDKSMVDPK-----KGTSLSDKLSPFSTHTLL 337

Db 66 -----SKGKATGRKAPLWLRAKFORLLFKLGCYIQNC 98

Qy 338 GQFFQCGTGWASWPLTLLIVLSVIPVVALAGLVTELTTPDVELWASPNQARSEKAPH 397

Db 99 GKF-----LVVGLLIFGAFVGLKAAANLETNVEELWVEVGGVRSRELYNT 143

Qy 398 DQHFGPFRTN-QVILTAPNRSSRYRYSLLGPKNFSGILDLDLLELE--LQERLRHL 454
Db 144 RQKIGEEAMFNQMIQTPKBEG-----ANVLTTEALLQHLSALQASRVHV 190
Qy 455 QWMSPEAQNRISLDQICVAP-----LNDPNTSLYDCCINSLSLQYQNNRTLLLLTA 505
Db 191 YMYN----RQWKLHLCKYKGBELITETGYMDQIIBYLYPCLITPLDCFEWAKLQSGTA 246
Qy 506 NOTLMGQ-----TSQVDKDHFLY-----CAN----- 527
Db 247 --YLLGKPLRWTFNDFPLEFLBELKKINYQVDSWEMLNKAHVGHGMDRPCLNADPDC 304
Qy 528 -----APL-----TFKQGT-----ALALSCMA 544
Db 305 PATAPNKNSTKPLDMALVNGGCHGLSRKYMHWQBELIVGGTVKNSTGKLVSAHALQTMF 364
Qy 545 DYGAVPFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRAFO 604
Db 365 QLMTEPKQMY--EHFKGYEY-----VSHINWNE---DKAAAILLEAMQRTTVEVVHQS 411
Qy 605 RRMAGMFQVTTAERSLDEINRTTAEDLPFATSVIVIFLYISLALGSYSWSRVMVDS 664
Db 412 AQNSTQKVLSPFT--TLDDILKSPDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
Qy 665 KATLGLGGVAVVVLGAVMAAGFFSYLGISSILVILQVVPFLVLSVGAONIFVLEYSORL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTVQLPFLALGVGVDDVFLAHAFST 525
Qy 725 PRPGEPREVHIGRALGRVAPSMMLCSLSEAI CFFLGAITMPAVRTFALTSGVLDPF 784
Db 526 QONKRIPFEDRTGECLKRTGASVALTSISNVTAFPMALIPIPALRAFSLQAAVVVVF 585
Qy 785 LLQMSAFVALLSLDSKROEASRLDVCCC-----VKPOEL-----PPG 822
Db 586 AMVLLIFAILSMDLRYRREDRLDIFCCFTSPCVSRVIOVEPOAVTTHDNTRYSPPPY 645
Qy 823 QGEGE----- 827
Db 646 SSHSPAHEQTIMQSTVQLRTEYDPTHVYVYTAEPSEISVQPVTVTQDTLSCQSPST 705
Qy 828 -----LLGFPQKAYAPFLHLLWITRGVLLFLALFGVSL 861
Db 706 SSTRDLLSQFSDSSLHCLPEPCTKWTLSFAEKHYAPLLPKAKVWVIFLGLLGVSL 765
Qy 862 YSMCHISVGLDQELALPKDSVLLDYFLNRYFVEGAPVYFVTTLGYNPFSSAGHNAICS 921
Db 766 YGTRKVRDGLDLDIVPRETREYDFIAAQKYFSF-----YNN----- 803
Qy 922 SAGCNPFSTQKIQYAT-----EPPEQSYLAI PASS-----WVDDRFIDWL----- 961
Db 804 -----YIVTQADYPNIQHLLYDLHRSFSNVKYVWLEENKQLPKMWLHYFDFMDLQGLQD 857
Qy 962 -----TPGS-----CCRLYISGPNKDKPCPSTVNSLNCNKMCSITMGS 1000
Db 858 AFSDSWETGKIMPNNYKNGSDGVLYAKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQHKYL-PWFLND-----RPNIKCPKGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYIYLTAWVSNDPVAYAAQANIRPHRPFWYHDK---ADYMPETRLRIPAAAP 968
Qy 1044 VLASRFMAYHKLKNSQDYTBALRAARELANITA-DLRKVPGTDPAPFVFPYTTITNVPY 1102
Db 969 IEYAQFPYINGLRDTSDFEAEIKVRTICSNYISLGLSSYPNG-----YFP-----LFW 1018
Qy 1103 EOYLTILPEGLPMLSLCLVPTFAVSCLLGLDLSGLNLLSIVMLVDTVGFMAWDIS 1162
Db 1019 EOYIGLRHWLLFLISVVLACTFLVCAVFLNFWTAGII-VMYALMTVELFGMGLIGK 1077
Qy 1163 YNAVSLINLVSAGMSVEFVSHITRSP--ALSTKPTWLERAKEATISMGSAVFAVAMTN 1220
Db 1078 LSAVPWVILIASVGVGVEFTVHVALAFLTAIGDN---RRVLALEHMFAPVLDG-AVST 1133

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:32:55 ; Search time 188 Seconds
(without alignments)
2960.365 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNRQF 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA_Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	US-10-621-758A-4	Sequence 4, Appli
2	6909	100.0	1332	US-10-663-208A-4	Sequence 4, Appli
3	6909	100.0	1332	US-10-646-301A-4	Sequence 4, Appli
4	6909	100.0	1332	US-10-736-769-4	Sequence 4, Appli
5	6896	99.8	1332	US-10-239-316-9	Sequence 9, Appli
6	6872.5	99.5	1359	US-10-621-758A-44	Sequence 44, Appl
7	6872.5	99.5	1359	US-10-663-208A-44	Sequence 44, Appl
8	6872.5	99.5	1359	US-10-646-301A-44	Sequence 44, Appl
9	6872.5	99.5	1359	US-10-736-769-44	Sequence 44, Appl
10	6536	94.6	1344	US-10-450-763-53052	Sequence 53052, A
11	5421.5	78.5	1331	US-10-621-758A-2	Sequence 2, Appli
12	5421.5	78.5	1331	US-10-663-208A-2	Sequence 2, Appli
13	5421.5	78.5	1331	US-10-646-301A-2	Sequence 2, Appli
14	5421.5	78.5	1331	US-10-736-769-2	Sequence 2, Appli
15	5407	78.3	1333	US-10-621-758A-12	Sequence 12, Appl
16	5407	78.3	1333	US-10-663-208A-12	Sequence 12, Appl
17	5407	78.3	1333	US-10-646-301A-12	Sequence 12, Appl
18	5407	78.3	1333	US-10-736-769-12	Sequence 12, Appl
19	4466	64.6	982	US-10-450-763-53050	Sequence 53050, A
20	2402.5	34.8	1278	US-10-208-731-2	Sequence 2, Appli
21	2402.5	34.8	1278	US-10-741-601-530	Sequence 530, App
22	2402.5	34.8	1278	US-10-741-600-1542	Sequence 1542, Ap
23	2402.5	34.8	1278	US-10-756-149-4924	Sequence 4924, Ap
24	2385	34.5	1319	US-10-208-731-4	Sequence 4, Appli
25	2262	32.7	1287	US-11-097-143-12003	Sequence 12003, A
26	1864.5	27.0	1223	US-11-097-143-2679	Sequence 2679, Ap
27	1329	19.2	1170	US-10-208-731-6	Sequence 6, Appli

28	1078	15.6	1296	4	US-10-208-731-9	Sequence 9, Appli
29	942	13.6	194	4	US-10-239-316-8	Sequence 8, Appli
30	900.5	13.0	492	4	US-10-424-599-211862	Sequence 211862,
31	859.5	12.4	541	4	US-10-424-599-189288	Sequence 189288,
32	651.5	9.4	1447	2	US-08-954-701A-19	Sequence 19, Appl
33	651.5	9.4	1447	3	US-09-898-533-5	Sequence 5, Appli
34	651.5	9.4	1447	3	US-09-754-032-19	Sequence 19, Appl
35	651.5	9.4	1447	4	US-10-421-446-19	Sequence 19, Appl
36	651.5	9.4	1447	4	US-10-791-844-6	Sequence 6, Appli
37	646.5	9.4	1434	2	US-08-954-701A-10	Sequence 10, Appl
38	646.5	9.4	1434	3	US-09-754-032-10	Sequence 10, Appl
39	646.5	9.4	1434	4	US-10-421-446-10	Sequence 10, Appl
40	628	9.1	1296	4	US-10-302-279-60	Sequence 60, Appl
41	598.5	8.7	933	4	US-10-415-934-3	Sequence 3, Appli
42	597	8.6	821	5	US-10-890-776A-4805	Sequence 4805, Ap
43	588	8.5	954	4	US-10-060-756A-3	Sequence 3, Appli
44	588	8.5	954	5	US-10-890-776A-3	Sequence 3, Appli
45	584	8.5	1203	3	US-09-990-046-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4

Query Match	100.0%	Score	6909;	DB	4;	Length	1332;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	1332;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;
QY	1	MAEAGLRGWLWALLRLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN	60				
Db	1	MAEAGLRGWLWALLRLAQSPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN	60				
QY	61	TPARKITGDHLTLLOKICPRLYTGPNTOACCSAKOLVLSLEASLSITKALLTRCPACSDNF	120				
Db	61	TPARKITGDHLTLLOKICPRLYTGPNTOACCSAKOLVLSLEASLSITKALLTRCPACSDNF	120				
QY	121	VNLHCHTCSNPQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSVDSCSRVRPAA	180				
Db	121	VNLHCHTCSNPQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSVDSCSRVRPAA	180				
QY	181	ATLVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNEGV	240				
Db	181	ATLVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDTFHLLRPGQAVGSGIQPLNEGV	240				
QY	241	ARCNESQGDVATCSQDCCAASCPAIRPQALDSTFYLQMPGSLVLIILCSVFAVVTI	300				
Db	241	ARCNESQGDVATCSQDCCAASCPAIRPQALDSTFYLQMPGSLVLIILCSVFAVVTI	300				
QY	301	LLVGRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLQOFFQCGWGTWVASWPLTILVLSV	360				
Db	301	LLVGRVAPARDKSNVDPKKGTSLSDKLSFSTHTLLQOFFQCGWGTWVASWPLTILVLSV	360				


```
Db 661 MVDKATLGLGGVAVVWLGAVMAAGFFSYLGRSSLVILQVVPFLVSLVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFPLGALTMPAVRTTALTSGLA 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFPLGALTMPAVRTTALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFFQKAYAPPL 840
Qy 841 LHWITRGVVLFLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPQSYLAIPASSWVDDFD 960
Db 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPQSYLAIPASSWVDDFD 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFVFPVYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
Db 1081 RKVPGTDPAFVFPVYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
Qy 1141 NLLSIVMLVDVTFGMALWDISYNAVSLINLSVAVGMSVEFVSHITRSPAI STKPTWLER 1200
Db 1141 NLLSIVMLVDVTFGMALWDISYNAVSLINLSVAVGMSVEFVSHITRSPAI STKPTWLER 1200
Qy 1201 AKEATISGSAFVAGVAMTNLPGIILVGLAKAQLIQIPFRLNLTILGLLHGLVFLPV 1260
Db 1201 AKEATISGSAFVAGVAMTNLPGIILVGLAKAQLIQIPFRLNLTILGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIVVNHSPGSIKGAGA 1320
Db 1261 ILSYVGPVDPNPALEQKRAEAAVAVMVASCPNHPSPRSTADNIVVNHSPGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
```

RESULT 3

US-10-646-301A-4

; Sequence 4, Application US/10646301A

; Publication No. US20040137467A1

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JB01603-K1-US

; CURRENT APPLICATION NUMBER: US/10/646,301A

; CURRENT FILING DATE: 2003-08-22

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 4

; LENGTH: 1332

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-646-301A-4

```
Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MABAGLGRWLLWALLRLAQSEPTTIIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLN 60
Db 1 MABAGLGRWLLWALLRLAQSEPTTIIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLN 60
Qy 61 TPARKITGDHLLILLOKI CRLPYTGNTQACCSAKOLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKI CRLPYTGNTQACCSAKOLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNPQSLFPI NVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHTCSNPQSLFPI NVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAGVTWCGVVGSAALCNAQRWLNFGQDGTGNGLA PLDITFHLLERQCAVSGIQLNBEV 240
Db 181 ATLAGVTWCGVVGSAALCNAQRWLNFGQDGTGNGLA PLDITFHLLERQCAVSGIQLNBEV 240
Qy 241 ARCNEQGDVVATCSQDCAASCPAIA RPAQALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNEQGDVVATCSQDCAASCPAIA RPAQALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
Qy 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLQOPFGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLQOPFGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDVELWSAPNSQASEKAFPHQHPGFPFRITNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQASEKAFPHQHPGFPFRITNQVILITAPNRSY 420
Qy 421 RYDLSLLGPKNPGSGLDLDLLELLEQLBRHLQWSPAEQARNISLODICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPGSGLDLDLLELLEQLBRHLQWSPAEQARNISLODICYAPLNPDNT 480
Qy 481 SLYDCINSLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGTALAL 540
Db 481 SLYDCINSLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGTALAL 540
Qy 541 SCMDYAGVAPVPPFLAIGYKGYSEALIMTFSLNYPAGDPRLAQNKLWEAFLEEM 600
Db 541 SCMDYAGVAPVPPFLAIGYKGYSEALIMTFSLNYPAGDPRLAQNKLWEAFLEEM 600
Qy 601 RAFORMAGMFOVTFPABRSLEDEINRTTAEDLPFATSYIIVFLYISLALGYSYSSWSRV 660
Db 601 RAFORMAGMFOVTFPABRSLEDEINRTTAEDLPFATSYIIVFLYISLALGYSYSSWSRV 660
Qy 661 MVDKATLGLGGVAVVWLGAVMAAGFFSYLGRSSLVILQVVPFLVSLVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVWLGAVMAAGFFSYLGRSSLVILQVVPFLVSLVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFPLGALTMPAVRTTALTSGLA 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFFPLGALTMPAVRTTALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFFQKAYAPPL 840
Qy 841 LHWITRGVVLFLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLFLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPQSYLAIPASSWVDDFD 960
Db 901 YFVTLGYNFSSEAGMNAICSSAGCNPFSTQKI QYATEFPQSYLAIPASSWVDDFD 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYL PWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQVLSRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
```

Db	1021	NIKCPKGLAAVSTSVNLTSDQVLASFMAHYKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAFEPVPTIINNVFYOVLTLPGELFMLSCLVPTFAVSCLLIGLDLRSGLL	1140
Db	1081	RKVPGTDPAFEPVPTIINNVFYOVLTLPGELFMLSCLVPTFAVSCLLIGLDLRSGLL	1140
Qy	1141	NLLSIVMLVDTVGFPMALWDISYNAVSLINLVASVGMSEVFSHITRSPAISTKPTWLER	1200
Db	1141	NLLSIVMLVDTVGFPMALWDISYNAVSLINLVASVGMSEVFSHITRSPAISTKPTWLER	1200
Qy	1201	AKEATISGSAVFAVAMTNLPGILVLGLAKAQLIQIIEFFRLNLLITLGLLHGLVFLPV	1260
Db	1201	AKEATISGSAVFAVAMTNLPGILVLGLAKAQLIQIIEFFRLNLLITLGLLHGLVFLPV	1260
Qy	1261	ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA	1320
Db	1261	ILSYVGPDPVPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA	1320
Qy	1321	ISNFLPNNGRQF 1332	
Db	1321	ISNFLPNNGRQF 1332	
RESULT 4			
US-10-736-769-4			
; Sequence 4, Application US/10736769			
; Publication No. US20040161838A1			
; GENERAL INFORMATION:			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JB01603-K3-US			
; CURRENT APPLICATION NUMBER: US/10/736,769			
; CURRENT FILING DATE: 2003-12-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; PRIOR APPLICATION NUMBER: 10/663,208			
; PRIOR FILING DATE: 2003-09-16			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 4			
; LENGTH: 1332			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-736-769-4			
Query Match 100.0%; Score 6909; DB 4; Length 1332;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	MAEAGLGLWLLALLRLAQSEPYTHIQPGYCAFYDECCKNPGLSGSLMTLSNVCSLSN	60
Db	1	MAEAGLGLWLLALLRLAQSEPYTHIQPGYCAFYDECCKNPGLSGSLMTLSNVCSLSN	60
Qy	61	TPARKITGDHLILLOKICPLRYTGPNTQACCSAKQLVLEASISITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICPLRYTGPNTQACCSAKQLVLEASISITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSNQSLFNVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDCSRVRVPA	180
Db	121	VNLHCHNTCSNQSLFNVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDCSRVRVPA	180
Qy	181	ATLAVGTWCGYVGSALCNAQRWLNFGDGTGNGLAPLDITPHLEPGQVSGIQPLNEGV	240
Db	181	ATLAVGTWCGYVGSALCNAQRWLNFGDGTGNGLAPLDITPHLEPGQVSGIQPLNEGV	240

Db 1321 ISNPLPNNGRQF 1332

RESULT 5

US-10-239-316-9

; Sequence 9, Application US/10239316

; Publication No. US20030125253A1

; GENERAL INFORMATION:

; APPLICANT: TANIYAMA, Yoshio

; APPLICANT: KITA, Shunbun

; APPLICANT: SATOMI, Tomoko Komiyama

; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And

; FILE REFERENCE: 2703USOP

; CURRENT APPLICATION NUMBER: US/10/239,316

; PRIOR FILING DATE: 2002-09-19

; PRIOR FILING DATE: 2001-03-22

; PRIOR FILING DATE: 2001-03-22

; PRIOR FILING DATE: 2000-03-24

; NUMBER OF SEQ ID NOS: 59

; SEQ ID NO 9

; LENGTH: 1332

; TYPE: PRT

; ORGANISM: Human

US-10-239-316-9

Query Match 99.8%; Score 6896; DB 4; Length 1332;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSBPYTHIQGYCAFYDECGKPELSGLMTLSNVSCLSN 60

Db 1 MAEAGLRGWLWALLRLAQSBPYTHIQGYCAFYDECGKPELSGLMTLSNVSCLSN 60

QY 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASITKALLTRCPACSNF 120

Db 61 TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVLSLEASITKALLTRCPACSNF 120

QY 121 VNLHCHTCSNQSLEFINTVRAQAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180

Db 121 VNLHCHTCSNQSLEFINTVRAQAGOLPAVVAEAFYQHSFAEQSYDSCSRVRPAA 180

QY 181 ATLVAGTMCVGYGSALCNAQRLNFGDGTGNGLAFLDITFHLLEPCQAVGSGIQPLNEGV 240

Db 181 ATLVAGTMCVGYGSALCNAQRLNFGDGTGNGLAFLDITFHLLEPCQAVGSGIQPLNEGV 240

QY 241 ARCNEQGDVATCSQDCAACPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

Db 241 ARCNEQGDVATCSQDCAACPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

QY 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360

Db 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360

QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFDQHFQFFRTNOVILTAPNRSY 420

Db 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFDQHFQFFRTNOVILTAPNRSY 420

QY 421 RYDLSLLGPKNPSGILDDLLELELQERLRLHQLWSPEAQRLNLSQDICYAPLNPDNT 480

Db 421 RYDLSLLGPKNPSGILDDLLELELQERLRLHQLWSPEAQRLNLSQDICYAPLNPDNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHPLYCANAPLTFCQGTALAL 540

Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHPLYCANAPLTFCQGTALAL 540

QY 541 SCWADYGAVFPFLATGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEAFLEEM 600

Db 541 SCWADYGAVFPFLATGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEAFLEEM 600

QY 601 RAFORMMAGMFQVTTFAERSLEDEINRTTAEDLPATSYIIVFLYISIALGYSWSRV 660

Db 601 RAFORMMAGMFQVTTFAERSLEDEINRTTAEDLPATSYIIVFLYISIALGYSWSRV 660

601 RAFORMMAGMFQVTTFAERSLEDEINRTTAEDLPATSYIIVFLYISIALGYSWSRV 660

661 MYDSKATLGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720

661 MYDSKATLGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLE 720

721 YQRLPRRGEPREVHIGRALGRVAPSMMLCSISEAICFFPLGALTMPAVRTTALTSLGLAV 780

721 YQRLPRRGEPREVHIGRALGRVAPSMMLCSISEAICFFPLGALTMPAVRTTALTSLGLAV 780

781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQELPPPGQGGSLGILGFPQKAYAPFL 840

781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQELPPPGQGGSLGILGFPQKAYAPFL 840

841 LHMWTRGVVLLIFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNLRYFEVGAPV 900

841 LHMWTRGVVLLIFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNLRYFEVGAPV 900

901 YFVTTLYGNFSEAGMNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960

901 YFVTTLYGNFSEAGMNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960

961 LTPSSCCRLYISGPNKDKFCBSTVNSLNCMLKNCMSITMGSVPSVEQEHKYLPMFLNDRP 1020

961 LTPSSCCRLYISGPNKDKFCBSTVNSLNCMLKNCMSITMGSVPSVEQEHKYLPMFLNDRP 1020

1021 NTKCPKGLAAYSTSVNLTSQGVLASRFMAHYKPLKNSQDYTEALRAARELAANITADL 1080

1021 NTKCPKGLAAYSTSVNLTSQGVLASRFMAHYKPLKNSQDYTEALRAARELAANITADL 1080

1081 RKVPGTDPAPFVPTTINNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140

1081 RKVPGTDPAPFVPTTINNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140

1141 NLLSTVMIILVDTVGPMAWLDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200

1141 NLLSTVMIILVDTVGPMAWLDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200

1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLILITLGLHLGLVFLPV 1260

1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIPFFRLNLILITLGLHLGLVFLPV 1260

1261 ILSYVGPDPVNPALALEQKRAEBAVAAMVAVASCPNHPSPRVSTADNIVNHSFEGSIKGAGA 1320

1261 ILSYVGPDPVNPALALEQKRAEBAVAAMVAVASCPNHPSPRVSTADNIVNHSFEGSIKGAGA 1320

1321 ISNPLPNNGRQF 1332

1321 ISNPLPNNGRQF 1332

RESULT 6

US-10-621-758A-44

; Sequence 44, Application US/10621758A

; Publication No. US20040093629A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Luquan

; APPLICANT: Altman, Scott W

; APPLICANT: Graziano, Michael

; TITLE OF INVENTION: NPC1L1

; FILE REFERENCE: JB01603-K-US

; CURRENT APPLICATION NUMBER: US/10/621,758A

; CURRENT FILING DATE: 2003-07-17

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 44

; LENGTH: 1359

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-621-758A-44

Query Match		99.5%;	Score 6872.5;	DB 4;	Length 1359;	
Best Local Similarity		97.9%;	Pred. No. 0;			
Matches 1330;		Conservative	0;	Mismatches	2;	Indels 27; Gaps 1;
Qy	1	MAEAGLRGWLWALLRLAQA	SEPYTTIHQPGYCAFYDECGK	NPELSGSLMTLSNV	CSLSN	60
Db	1	MAEAGLRGWLWALLRLAQA	SEPYTTIHQPGYCAFYDECGK	NPELSGSLMTLSNV	CSLSN	60
Qy	61	TPARKITGDHLLLOKICPRL	YTGNTQACCSAKQLVSEAS	LSTIKALLTRCPAC	SDNF	120
Db	61	TPARKITGDHLLLOKICPRL	YTGNTQACCSAKQLVSEAS	LSTIKALLTRCPAC	SDNF	120
Qy	121	VNLHCHNTCSPNQSLFIN	TRVAQLGAGQLPAVAYEAF	YQHSFABQSYDSC	SRVRPAA	180
Db	121	VNLHCHNTCSPNQSLFIN	TRVAQLGAGQLPAVAYEAF	YQHSFABQSYDSC	SRVRPAA	180
Qy	181	ATLAVGTMCVGYGSA	LCNAQRWLNFGQD	TGNGLAPLDITFHL	LEPGQVSGIQPLNE	GV 240
Db	181	ATLAVGTMCVGYGSA	LCNAQRWLNFGQD	TGNGLAPLDITFHL	LEPGQVSGIQPLNE	GV 240
Qy	241	ARCNESQDDVATCS	QDCDCAASCPA	ITARPQALDSTF	VLGQMPGSLVLI	ILICSVFAVVTI 300
Db	241	ARCNESQDDVATCS	QDCDCAASCPA	ITARPQALDSTF	VLGQMPGSLVLI	ILICSVFAVVTI 300
Qy	301	LLVGFVAPARDKSM	VDPKGTSLSDXLSF	STHTLLGQFFQ	CGMTWASWPLT	LVLVSV 360
Db	301	LLVGFVAPARDKSM	VDPKGTSLSDXLSF	STHTLLGQFFQ	CGMTWASWPLT	LVLVSV 360
Qy	361	IPVVALAAGLVF	TELTDPVELWAS	PNQARSEKAFH	DQHPFRTN	QVILITAPNRSY 420
Db	361	IPVVALAAGLVF	TELTDPVELWAS	PNQARSEKAFH	DQHPFRTN	QVILITAPNRSY 420
Qy	421	RYDSLLGPKNF	SGIILDLLOLLLE	LQERLRLH	QVWSP	EAQRNLSLODICVAPLNPDNT 480
Db	421	RYDSLLGPKNF	SGIILDLLOLLLE	LQERLRLH	QVWSP	EAQRNLSLODICVAPLNPDNT 480
Qy	481	SLYDCCINS	LLQYFQNNRTLL	LTANTQ	LMGQTSQVD	WDKHFYCANAPLTKDGTALAL 540
Db	481	SLYDCCINS	LLQYFQNNRTLL	LTANTQ	LMGQTSQVD	WDKHFYCANAPLTKDGTALAL 540
Qy	541	SWADYCAP	VPFLAIGYKGDY	SEAEALIM	TFSLNNYPAG	DPRLAQAKLWEEAFLEEM 600
Db	541	SWADYCAP	VPFLAIGYKGDY	SEAEALIM	TFSLNNYPAG	DPRLAQAKLWEEAFLEEM 600
Qy	601	RAFORMAGM	FQVTFTAERS	LEDEINRTT	AEIDLPIFATS	YIVITFLYISIALGYSWSRV 660
Db	601	RAFORMAGM	FQVTFTAERS	LEDEINRTT	AEIDLPIFATS	YIVITFLYISIALGYSWSRV 660
Qy	661	MVDSKAT	GLGGVAVVL	GAVMAANG	PFYSYLGIR	SSVLQVVPFLVSGADNIFLVL 720
Db	661	MVDSKAT	GLGGVAVVL	GAVMAANG	PFYSYLGIR	SSVLQVVPFLVSGADNIFLVL 720
Qy	721	YQRLPRP	GPPEVHIGRAL	GRVAPSM	LLCSLS	EACFFILGALTPMPAVRTFALTSLAV 780
Db	721	YQRLPRP	GPPEVHIGRAL	GRVAPSM	LLCSLS	EACFFILGALTPMPAVRTFALTSLAV 780
Qy	781	ILDFLQMS	AFVALLSLDS	KRQEASRL	DVCCVKPQ	ELPPPGQEGLLLFFQKAYAPFL 840
Db	781	ILDFLQMS	AFVALLSLDS	KRQEASRL	DVCCVKPQ	ELPPPGQEGLLLFFQKAYAPFL 840
Qy	841	LHWITRG	VVLLFLAL	FGVSLYS	WCHISV	GLDQELAPKOSYLDYFLNRYFEVGPV 900
Db	841	LHWITRG	VVLLFLAL	FGVSLYS	WCHISV	GLDQELAPKOSYLDYFLNRYFEVGPV 900
Qy	901	YFVTT	LGYNFSEAG	WNAICSS	SAGCNF	SFTQKIQYATEFPEQSYLAIPASSWDDFIDW 960
Db	901	YFVTT	LGYNFSEAG	WNAICSS	SAGCNF	SFTQKIQYATEFPEQSYLAIPASSWDDFIDW 960
Qy	961	LTPSSC	CRXYISGPN	KDPCSTV	NSLNC	KNCMSITWGSVRPSVEQPHKYLPHFLNDRP 1020
Db	961	LTPSSC	CRXYISGPN	KDPCSTV	NSLNC	KNCMSITWGSVRPSVEQPHKYLPHFLNDRP 1020
RESULT 7						
US-10-663-208A-44						
; Sequence 44, Application US/10663208A						
; Publication No. US20040132058A1						
; GENERAL INFORMATION:						
; APPLICANT: Wang, Luquan						
; APPLICANT: Graziano, Michael						
; APPLICANT: Murgolo, Nick						
; TITLE OF INVENTION: NFK1L1 (NPC3) AND METHODS OF USE THEREOF						
; FILE REFERENCE: JB01603K2 US						
; CURRENT APPLICATION NUMBER: US/10/663,208A						
; CURRENT FILING DATE: 2003-09-16						
; PRIOR APPLICATION NUMBER: 60/397,442						
; PRIOR FILING DATE: 2002-07-19						
; PRIOR APPLICATION NUMBER: 10/621,758						
; PRIOR FILING DATE: 2003-07-17						
; PRIOR APPLICATION NUMBER: 10/646,301						
; PRIOR FILING DATE: 2003-08-22						
; NUMBER OF SEQ ID NOS: 50						
; SOFTWARE: PatentIn version 3.1						
; SEQ ID NO 44						
; LENGTH: 1359						
; TYPE: PRT						
; ORGANISM: Homo sapiens						
US-10-663-208A-44						
Query Match		99.5%;	Score 6872.5;	DB 4;	Length 1359;	
Best Local Similarity		97.9%;	Pred. No. 0;			
Matches 1330;		Conservative	0;	Mismatches	2;	Indels 27; Gaps 1;
Qy	1	MAEAGLRGWLWALLRLAQA	SEPYTTIHQPGYCAFYDECGK	NPELSGSLMTLSNV	CSLSN	60
Db	1	MAEAGLRGWLWALLRLAQA	SEPYTTIHQPGYCAFYDECGK	NPELSGSLMTLSNV	CSLSN	60
Qy	61	TPARKITGDHLLLOKICPRL	YTGNTQACCSAKQLVSEAS	LSTIKALLTRCPAC	SDNF	120
Db	61	TPARKITGDHLLLOKICPRL	YTGNTQACCSAKQLVSEAS	LSTIKALLTRCPAC	SDNF	120
Qy	121	VNLHCHNTCSPNQSLFIN	TRVAQLGAGQLPAVAYEAF	YQHSFABQSYDSC	SRVRPAA	180
Db	121	VNLHCHNTCSPNQSLFIN	TRVAQLGAGQLPAVAYEAF	YQHSFABQSYDSC	SRVRPAA	180
Qy	181	ATLAVGTMCVGYGSA	LCNAQRWLNFGQD	TGNGLAPLDITFHL	LEPGQVSGIQPLNE	GV 240
Db	181	ATLAVGTMCVGYGSA	LCNAQRWLNFGQD	TGNGLAPLDITFHL	LEPGQVSGIQPLNE	GV 240
Qy	241	ARCNESQDDVATCS	QDCDCAASCPA	ITARPQALDSTF	VLGQMPGSLVLI	ILICSVFAVVTI 300
Db	241	ARCNESQDDVATCS	QDCDCAASCPA	ITARPQALDSTF	VLGQMPGSLVLI	ILICSVFAVVTI 300
Qy	301	LLVGFVAPARDKSM	VDPKGTSLSDXLSF	STHTLLGQFFQ	CGMTWASWPLT	LVLVSV 360
Db	301	LLVGFVAPARDKSM	VDPKGTSLSDXLSF	STHTLLGQFFQ	CGMTWASWPLT	LVLVSV 360
Qy	361	IPVVALAAGLVF	TELTDPVELWAS	PNQARSEKAFH	DQHPFRTN	QVILITAPNRSY 420
Db	361	IPVVALAAGLVF	TELTDPVELWAS	PNQARSEKAFH	DQHPFRTN	QVILITAPNRSY 420
Qy	421	RYDSLLGPKNF	SGIILDLLOLLLE	LQERLRLH	QVWSP	EAQRNLSLODICVAPLNPDNT 480
Db	421	RYDSLLGPKNF	SGIILDLLOLLLE	LQERLRLH	QVWSP	EAQRNLSLODICVAPLNPDNT 480
Qy	481	SLYDCCINS	LLQYFQNNRTLL	LTANTQ	LMGQTSQVD	WDKHFYCANAPLTKDGTALAL 540
Db	481	SLYDCCINS	LLQYFQNNRTLL	LTANTQ	LMGQTSQVD	WDKHFYCANAPLTKDGTALAL 540
Qy	541	SWADYCAP	VPFLAIGYKGDY	SEAEALIM	TFSLNNYPAG	DPRLAQAKLWEEAFLEEM 600
Db	541	SWADYCAP	VPFLAIGYKGDY	SEAEALIM	TFSLNNYPAG	DPRLAQAKLWEEAFLEEM 600
Qy	601	RAFORMAGM	FQVTFTAERS	LEDEINRTT	AEIDLPIFATS	YIVITFLYISIALGYSWSRV 660
Db	601	RAFORMAGM	FQVTFTAERS	LEDEINRTT	AEIDLPIFATS	YIVITFLYISIALGYSWSRV 660
Qy	661	MVDSKAT	GLGGVAVVL	GAVMAANG	PFYSYLGIR	SSVLQVVPFLVSGADNIFLVL 720
Db	661	MVDSKAT	GLGGVAVVL	GAVMAANG	PFYSYLGIR	SSVLQVVPFLVSGADNIFLVL 720
Qy	721	YQRLPRP	GPPEVHIGRAL	GRVAPSM	LLCSLS	EACFFILGALTPMPAVRTFALTSLAV 780
Db	721	YQRLPRP	GPPEVHIGRAL	GRVAPSM	LLCSLS	EACFFILGALTPMPAVRTFALTSLAV 780
Qy	781	ILDFLQMS	AFVALLSLDS	KRQEASRL	DVCCVKPQ	ELPPPGQEGLLLFFQKAYAPFL 840
Db	781	ILDFLQMS	AFVALLSLDS	KRQEASRL	DVCCVKPQ	ELPPPGQEGLLLFFQKAYAPFL 840
Qy	841	LHWITRG	VVLLFLAL	FGVSLYS	WCHISV	GLDQELAPKOSYLDYFLNRYFEVGPV 900
Db	841	LHWITRG	VVLLFLAL	FGVSLYS	WCHISV	GLDQELAPKOSYLDYFLNRYFEVGPV 900
Qy	901	YFVTT	LGYNFSEAG	WNAICSS	SAGCNF	SFTQKIQYATEFPEQSYLAIPASSWDDFIDW 960
Db	901	YFVTT	LGYNFSEAG	WNAICSS	SAGCNF	SFTQKIQYATEFPEQSYLAIPASSWDDFIDW 960
Qy	961	LTPSSC	CRXYISGPN	KDPCSTV	NSLNC	KNCMSITWGSVRPSVEQPHKYLPHFLNDRP 1020
Db	961	LTPSSC	CRXYISGPN	KDPCSTV	NSLNC	KNCMSITWGSVRPSVEQPHKYLPHFLNDRP 1020

Db 241 ARCNESQDDVATCSCQDCAASCPAIPQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELEQLERLHLQVWSPEAQNRNLSQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELEQLERLHLQVWSPEAQNRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
QY 541 SCADYGAPVFPPLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCADYGAPVFPPLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
QY 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPATSYIVIFLYISALGYSYSSWSRV 660
Db 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPATSYIVIFLYISALGYSYSSWSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAWGPFSYLGIRSSLVILQVVPFLVLSVGADNIPFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAWGPFSYLGIRSSLVILQVVPFLVLSVGADNIPFVLE 720
QY 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKVQOELPPGQEGELLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKROEASRLDVCCCKVQOELPPGQEGELLGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALPGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNRYFEVGAVP 900
Db 841 LHWITRGVLLFLALPGVSLYSMCHI SVGLDQELALPKDSYLLDYFLPLNRYFEVGAVP 900
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSTQKIQVATEFPFQSYLAIPASSWVDDFD 960
Db 901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSTQKIQVATEFPFQSYLAIPASSWVDDFD 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVTNSLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAVSTSVNLTSDGOVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAVSTSVNLTSDGOVLDTVAILSPRLEYSTISAHNCNLYLLDSASRPMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAPFVFPYTTINVFYEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAPFVFPYTTINVFYEQVLTILPEGL 1140
QY 1114 FMLSLCLVPTFAVSCLLGLDLSRGLNLLSVIMLVDPVGMALWDISYNAVSLINLVS 1173
Db 1141 FMLSLCLVPTFAVSCLLGLDLSRGLNLLSVIMLVDPVGMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHLTRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233
Db 1201 AVGMSVEFVSHLTRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIFPFRNLNLTILGLLHGLVFLPVILSYVGPVNPALAEQKRAEAAVAVWVASC 1293
Db 1261 LIQIFPFRNLNLTILGLLHGLVFLPVILSYVGPVNPALAEQKRAEAAVAVWVASC 1320
QY 1294 NHPSRVSTADNTIYNHSPFGSIKGAISNFIPLNNGRQF 1332

Db 1321 NHPSRVSTADNTIYNHSPFGSIKGAISNFIPLNNGRQF 1359
RESULT 8
US-10-646-301A-44
; Sequence 44, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Luquan
; APPLICANT: Altman, Scott W
; APPLICANT: Graziano, Michael
; APPLICANT: Murgo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-646-301A-44
Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 MAEAGLRGWLWALLRLAQSSEPYTHIOPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSSEPYTHIOPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTPGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILQKICPRLYTPGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSILFINVTRVAQIAGOLPAVAYEAPYQHSFARQSDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSILFINVTRVAQIAGOLPAVAYEAPYQHSFARQSDSCSRVRPAA 180
QY 181 ATLAAGTMCVGVGSALCNAQRLNFGDGTGNGLAPLIDITFHLLPEQAVGSGIQIPNEGV 240
Db 181 ATLAAGTMCVGVGSALCNAQRLNFGDGTGNGLAPLIDITFHLLPEQAVGSGIQIPNEGV 240
QY 241 ARCNESQDDVATCSCQDCAASCPAIPQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNESQDDVATCSCQDCAASCPAIPQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELEQLERLHLQVWSPEAQNRNLSQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELEQLERLHLQVWSPEAQNRNLSQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTPKDGFTALAL 540
QY 541 SCADYGAPVFPPLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCADYGAPVFPPLAIGGYKGDYSEBALIMTFSLNNYPAGDPRLAQAQKWEAFLEEM 600
QY 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPATSYIVIFLYISALGYSYSSWSRV 660

Db 601 RAQRMAGMFOVTFMAERSLEDEINRTTADLPFATSYIVIFLYISLALGSSSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPQOGEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPQOGEGLLGFFQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNVNFSTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNVNFSTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPTVNSLNCCKMCSITWGSVRPSVEQFHKYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPTVNSLNCCKMCSITWGSVRPSVEQFHKYLPWFLNDRP 1020
Qy 1021 NIKCPKGGIAAYSTSVNLTSDQOVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGIAAYSTSVNLTSDQOVLDTVAILSPRLEYSCTTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVPPYTIITNVFVEQVLTILPSEL 1113
Db 1081 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVPPYTIITNVFVEQVLTILPSEL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSTVMILVDVTGFMALWDISYNAVSLINLV 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSTVMILVDVTGFMALWISYNAVSLINLV 1200
Qy 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGIILVLGAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGIILVLGAKAQ 1260
Qy 1234 LIQIFPFRNLTLTLGLHLGVFLPVILSYVGPVNPALAEOKBAEAAVAAMVWASCP 1293
Db 1261 LIQIFPFRNLTLTLGLHLGVFLPVILSYVGPVNPALAEOKBAEAAVAAMVWASCP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFGSIKAGAGAINFNLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAGAINFNLPNNGRQF 1359

RESULT 9
US-10-736-769-44
; Sequence 44, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10736, 769
; PRIOR FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16

; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLAGWLLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPGLSGLMTLSNVCLSN 60
Db 1 MAEAGLAGWLLWALLRLAQSEPYTTIHQPGYCAFYDECCKNPGLSGLMTLSNVCLSN 60
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLITKALLTRCPACSNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGPNTOACCSAKQLVSLASLITKALLTRCPACSNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAGVTCGYGYSALCNQAWLNFGQDGTGNGLAPLDITFHLBPQAVGSGIQLNEGV 240
Db 181 ATLAGVTCGYGYSALCNQAWLNFGQDGTGNGLAPLDITFHLBPQAVGSGIQLNEGV 240
Qy 241 ARCHESQDDVATCSCQDCAASCALARPQALDSTFYLGQMPGSLVLIILICSVPAVVTI 300
Db 241 ARCHESQDDVATCSCQDCAASCALARPQALDSTFYLGQMPGSLVLIILICSVPAVVTI 300
Qy 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFTHTLGLQFFQGGTGWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFTHTLGLQFFQGGTGWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTTELTPDVELWAPNSQARSEKAFHQHFGPPFRTHQVILTAENRSY 420
Db 361 IPVVALAAGLVFTTELTPDVELWAPNSQARSEKAFHQHFGPPFRTHQVILTAENRSY 420
Qy 421 RYDLSLLIGPKRPSGTLDDIALELLEQLERHLQVWSPQARNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLIGPKRPSGTLDDIALELLEQLERHLQVWSPQARNISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGFTAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGFTAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLBEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLBEM 600
Qy 601 RAQRMAGMFOVTFMAERSLEDEINRTTADLPFATSYIVIFLYISLALGSSSRV 660
Db 601 RAQRMAGMFOVTFMAERSLEDEINRTTADLPFATSYIVIFLYISLALGSSSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPQOGEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVPOELPPQOGEGLLGFFQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNVNFSTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960


```
Db 901 YFTTIGYFSSAGMNAICSSAGCNFSFTQKIYATFPPQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVEQFKYLPWFNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVEQFKYLPWFNDRP 1020
Qy 1021 NIKCPKGGAAVSTSVNLTSQGV-----ASRFMAYH 1053
Db 1021 NIKCPKGGAAVSTSVNLTSQGVLDTVAILSPRLYSGTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVGGTDPAPFPYPTITNVFYEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVGGTDPAPFPYPTITNVFYEQVLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCILLGLDRLSGLNLLSIWMLVDTVGFMAWLDISYNAVSLNLYS 1173
Db 1141 FMLSCLVPTFAVSCILLGLDRLSGLNLLSIWMLVDTVGFMAWLGISYNAVSLNLYS 1200
Qy 1174 AVGMSVEFVSHITRSPASTKPTWLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPASTKPTWLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQ 1260
Qy 1234 LIQIFPFRNLAIITLGLHGLVFLVILSYVGPDPNPALEQKRAEBAVAAMVASCP 1293
Db 1261 LIQIFPFRNLAIITLGLHGLVFLVILSYVGPDPNPALEQKRAEBAVAAMVASCP 1320
Qy 1294 NHPSRVSTADNIVNHSFSGIKGAGAINFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIVNHSFSGIKGAGAINFLPNNGRQF 1359

RESULT 10
US-10-450-763-53052
; Sequence 53052, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)..(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052
```

```
Query Match 94.6%; Score 6536; DB 5; Length 1344;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;
```

```
Qy 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFYDECGKPELSGSLMTLSNVSCLSN 60
```

```
Qy 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSLPFINTRVAQLGAGQLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSLPFINTRVAQLGAGQLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
Qy 181 ATLAVGTMCVGVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQAVCGSGIQLNEGV 240
Db 181 ATLAVGTMCVGVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQAVCGSGIQLNEGV 240
Qy 241 ARCNSQGDVATCSCQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILCSVFVV 298
Db 241 ARCNSQGDVATCSCQDCAASCP--AIARPOALDSTFYLGQMPGSLVLIILCSVFVV 298
Qy 299 TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTILGQPGQGTWVASHPLILVL 358
Db 299 TILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTILGQPGQGTWVASHPLILVL 358
Qy 359 SVIPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGPPFRFRTNOVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGPPFRFRTNOVILTAPNRS 418
Qy 419 SYRYSLLGPKNPFSGIILDLDLLELLELLELLELLELLELLELLELLELLELLELLEL 478
Db 419 SYRYSLLGPKNPFSGIILDLDLLELLELLELLELLELLELLELLELLELLELLELLEL 478
Qy 479 NTSLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFYCANAPLTFFKGTAL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFYCANAPLTFFKGTAL 538
Qy 539 ALSCHADYGAIVPFLAIGYGGKDYSEAEALIMTFLSNYPAGDPRLAQKLEWESAFLE 598
Db 539 ALSCHADYGAIVPFLAIGYGGKDYSEAEALIMTFLSNYPAGDPRLAQKLEWESAFLE 598
Qy 599 EMRAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPFATSVIYIYISLALGYSYSSWS 658
Db 599 EMRAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPFATSVIYIYISLALGYSYSSWS 658
Qy 659 RVWVDSKATLGLGGVAVVLGAVMAAGPPSYIGIRSSVILQVVPFLVSVGADNIFIFV 718
Db 659 RVWVDSKATLGLGGVAVVLGAVMAAGPPSYIGIRSSVILQVVPFLVSVGADNIFIFV 718
Qy 719 LEYQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAICFELGALTMPAVTTEALTSG 778
Db 719 LEYQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAICFELGALTMPAVTTEALTSG 778
Qy 779 AVILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGILLGFFOKAYAP 838
Db 779 AVILDFLLQMSAFVALLSLDSKQBSRLDVCCVKPQELPPPGQEGILLGFFOKAYAP 838
Qy 839 FLLHWITRGV-----LFLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF 894
Db 839 FLLHWITRGV-----LFLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF 894
Qy 895 EVGAPVYFVTTLYGYNFSSEAGNNAICSSAGCNFSFTQKIYATFPPQSYLAIPASSWV 954
Db 895 EVGAPVYFVTTLYGYNFSSEAGNNAICSSAGCNFSFTQKIYATFPPQSYLAIPASSWV 954
Qy 955 DDFIDWLTSPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVEQFKYLPW 1014
Db 955 DDFIDWLTSPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSTMGSVRPSVEQFKYLPW 1014
Qy 1015 FLNDRPNTKCPKGGAAVSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAA 1074
Db 1015 FLNDRPNTKCPKGGAAVSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAA 1074
Qy 1075 NITADLRKVPCTDPAFEVPPYPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCILLGLD 1134
Db 1075 NITADLRKVPCTDPAFEVPPYPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCILLGLD 1134
Qy 1135 LRSGLNLLSIWMLVDTVGFMAWLDISYNAVSLNLYSAGVMSVEFVSHITRSPASTK 1194
```

```
Db 1139 LRSGLNLSIVMLVDTVGFMAWGISYNAVSLINLSVAVGMSVEFVSHITRSPALSTK 1198
Qy 1195 PTLWERAKEATISMSAVFAGVAMTNLPGILVLGLAKAQLQIIFFRNLNLTLLGLLHG 1254
Db 1199 PTLWERAKEATISMSAVFAGVAMTNLPGILVLGLAKAQLQIIFFRNLNLTLLGLLHG 1258
Qy 1255 LVFLPVILSYGPDVNPALALEQKRAEA-----VAAMVAVASCPNHPHRSVST 1301
Db 1259 LVFLPVILSYGPDVNPALALEQKRAEGRGGSGHGPQAQITPPESPQLTTSMT 1313

RESULT 11
US-10-621-758A-2
; Sequence 2, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2

Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
Db 1 MAEAWL-GWLLWALLLSAQQELYTPKHEAGVCTFEYECGKPELSGGLTSLSNVCSLSN 59
Qy 61 TPARKITGHLILLOKICRLYTGPNTO-ACCSAKOLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGHLALLOKICRLYNGPNTTFACCSKQLLSLESSMSITKALLTRCPACSDN 119
Qy 120 FVNLHCHNTCSNQSIFINVTVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPA 179
Db 120 FVSLHCHNTCSPOSLFINVTVRVERGAGEPPAVAYEAFYQHSFAEKAYESCVRIPA 179
Qy 180 AATLAVGTCMGVYGSALCNAQRWLNFGDGTGNGLAPLDTIFHLLPEQAVGSGIQLNPG 239
Db 180 AASLAVGSMGVYGSALCNAQRWLNFGDGTGNGLAPLDTIFHLLPEQALPDGIQLNGK 239
Qy 240 VARNESQGDVATCSQDCAACPAIARQALDSTFYLGQMPGSLVLIILCSFVAVT 299
Db 240 IAPCNESQGDSDSAVCSQDCAACPVIPPEALRPSFYMGPMGWLALIIIFAVFVLVS 299
Qy 300 ILVGFVAVARADSKMVDPKGTSLSKLSFSHTLLGQFGQGWGTVWASWPLTILVLS 359
Db 300 AVLRLRVNRRNKNKAEQGQAPKLPKHKLSPHILGRFQWNGTRVWASWPLTVLALS 359
Qy 360 VIPVALAAGLFTFELTTPVELWSAPNSQARSEKAFHQHFGPPFRFTNQVILTAENRSS 419
Db 360 FIVVIALAAGLFTFELTTPVELWSAPNSQARSEKSFHDSHFPGPPFRFTNQIFVTARNRSS 419
Qy 420 YRYDSLLGPKNFSGILDLDLLELLELQERLHLQWSPQAQRNLSLDICVAPLNPDN 479
Db 420 YKYSLLLGSKNFSGILSLDFLELLELQERLHLQWSPQAQRNLSLDICVAPLNPYN 479
Qy 480 TSLYDCCINSLSLQVFQNNRTLLLTANQTLNGQTSQVMDKDHFLYCANAPLTFKDGTSLA 539
```

```
Db 480 TSLDCCVNSLLQVFQNNRTLLMTANQTLNGQTSQVMDKDHFLYCANAPLTFKDGTSLA 539
Qy 540 LSCWADYGAPVFPFLAIGGYKGYDSEAEALIMTFESLNYPAGDPRLAQAKLWEEAFLEE 599
Db 540 LSCWADYGAPVFPFLAIGGYKGYDSEAEALIMTFESLNYPAGDPRLAQAKLWEEAFLEE 599
Qy 600 MRAPFORMAGMPOVTFPERSLEDEINRTTAEDLPFATSYIVIFLYISLALSGYSWSR 659
Db 600 MESFORNTSKQVAFSAERSLEDEINRTTIQDLPPFAVSIIIVFLYISLALSGYSRCSR 659
Qy 660 VMVDSKATLGLGGVAVVILGAVMAAMGFYSYLGTRSSLVILQVPPFLVLSGVADNIFIFVL 719
Db 660 VAVESKATLGLGGVIVVILGAVLAAMGFYSYLGVPSSLVIIQVWPFLVAVGADNIFIFVL 719
Qy 720 EYQRLPRRGEPRVHIGRALGRVAPDSMLCSSEALCFFLGALTMPAVRTFALTSGLA 779
Db 720 EYQRLPRMGEQREAHIGRTIGSVAPSMLLCSSEALCFFLGALTMPAVRTFALTSGLA 779
Qy 780 VILDFLLQMSAFVALLSDSKRQESRLDVCCKVPOELPPPQCGEGLLGFFOKAYAPF 839
Db 780 IILDFLLQMTAFVALLSDSKRQESRDPVLCCTFTRKLPPEKKEGGLLRFRKRIYAPF 839
Qy 840 LLHWITRGVLLFLFALFGVSLYSMCHISVGLDDELALPKDSYLLDYLFLNRYFVETGAP 899
Db 840 LLHRFIRPVVMLFLTLFGANLYLMCNINVLGDELALPKDSYLLDYLFLNRYLEVGP 899
Qy 900 VYFVTTLYGPFSSAGNATCSSAGCNFSFTQIOYATBPEQSYLAIPASSWVDDPID 959
Db 900 VYFVTTGFGNFSAGMNATCSSAGCKSFSUTQIOYASEFPDQSYVNAIASSWVDDPID 959
Qy 960 WLTP-SSCCRLYISGPNKDFCPSVTNSLNCNCKMSITMGSVRPSVEQFHXYLPWFND 1018
Db 960 WLTPSSCCRLYIRGPHKDFCPSVTDTSFNCLKNCNRTLGPVPTAEQFHXYLPWFND 1019
Qy 1019 RPNKCPKGLAAAYSTSVNLSDQVLAASRMAYHKLKNSQDTEALRAARELAANITA 1078
Db 1020 PENIRCPKGLAAAYRTSVNLSSDQVLAASQPMAYHKLKNSQDTEALRASRLAANITA 1079
Qy 1079 DLKRVGTDPAFEPFVPTITNVFEYOYLTILPEGLFMLSCLVPTFAVSCILGLDLSRG 1138
Db 1080 DLKRVGTDPAFEPFVPTITNVFEYOYLTILPEGLFMLSCLVPTFAVSCILGLDLSRG 1139
Qy 1139 LILNLSIVMLVDTVGFMAWMDISYNAVSLINLSVAVGMSVEFVSHITRSPALSTKPTWL 1198
Db 1140 LILNLSIIMLVDTIGLMAVWGISYNAVSLINLSVAVGMSVEFVSHITRSPALSTKPTWL 1199
Qy 1199 ERAKATISMSAVFAGVAMTNLPGILVLGLAKAQLQIIFFRNLNLTLLGLLGLVFL 1258
Db 1200 ERAKATISMSAVFAGVAMTNLPGILVLGLAKAQLQIIFFRNLNLTLLGLLGLVFL 1259
Qy 1259 PVILSYGPDVNPALALEQKRAEAFAVAVASCPNHPHRSVSTADNIYVNHSEFSGI-KG 1317
Db 1260 PVILSYGPDVNPALALEQKRAEAFAVAVASCPNHPHRSVSTADNIYVNHSEFSGI-KG 1317
Qy 1318 AGAISNLFNPNGRQF 1332
Db 1317 ANAARSLPKSQKF 1331

RESULT 12
US-10-663-208A-2
; Sequence 2, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
```

```
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-663-208A-2

Query Match
Best Local Similarity 78.5%; Score 5421.5; DB 4; Length 1331;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 59

QY 61 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVYAYEAFYQHSFAQSYDCSRVRVPA 179
Db 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVYAYEAFYQHSFAQSYDCSRVRVPA 179

QY 180 RATLAVGTMCVYGSALCNAQWLNFGQDTGNGLAPLDTITFLLRPGQAVGSIQPLNEG 239
Db 180 RATLAVGTMCVYGSALCNAQWLNFGQDTGNGLAPLDTITFLLRPGQAVGSIQPLNEG 239

QY 240 VARNCSQDDVATSCDCAACPAIARPOALDSTFYLGMPGSLVLIILCSYFAVVT 299
Db 240 VARNCSQDDVATSCDCAACPAIARPOALDSTFYLGMPGSLVLIILCSYFAVVT 299

QY 300 ILLVGRVAPARDKSNVDPKGTSLSDKLSPTSTHLLGQPGQGWVWASWPLTILVLS 359
Db 300 ILLVGRVAPARDKSNVDPKGTSLSDKLSPTSTHLLGQPGQGWVWASWPLTILVLS 359

QY 360 VTPVVALAAGLVFTLITDTPVLSNAPSOARSEKAFHQHGGPFRFTNOVLTAPNRSS 419
Db 360 VTPVVALAAGLVFTLITDTPVLSNAPSOARSEKAFHQHGGPFRFTNOVLTAPNRSS 419

QY 420 KYRYSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSIQDICVAPLNPDN 479
Db 420 KYRYSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSIQDICVAPLNPDN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDGTLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTFSLNYPAGDPRRLAQAKLWEAPLEE 599
Db 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTFSLNYPAGDPRRLAQAKLWEAPLEE 599

QY 600 MRAPFORMAGMFQVFTFAERSLEDEINRTTADLPATSYIVIFLYISLALGYSYSSWR 659
Db 600 MRAPFORMAGMFQVFTFAERSLEDEINRTTADLPATSYIVIFLYISLALGYSYSSWR 659

QY 660 VVMSDKATLGLGGVAVLGVANMAAGFFSYLGRSLVLQVVPFLVLSVGADNIFIFVL 719
Db 660 VVMSDKATLGLGGVAVLGVANMAAGFFSYLGRSLVLQVVPFLVLSVGADNIFIFVL 719

QY 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFLGALTPMAVTRTFALTSLG 779
Db 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFLGALTPMAVTRTFALTSLG 779

QY 780 VILDFLLQMSAFVALLSDSKQEARLDVCCVQPKQLPPQCGEGLLGFFQKAYAPF 839
Db 780 VILDFLLQMSAFVALLSDSKQEARLDVCCVQPKQLPPQCGEGLLGFFQKAYAPF 839

QY 840 LLHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFPLNRYPRVGAP 899
Db 840 LLHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFPLNRYPRVGAP 899

QY 900 VYFVTTLGYNPSSEAGMNAICSSACGNPSPFKIQYATEPPEQSYLAIPASSWVDDFID 959
Db 900 VYFVTTLGYNPSSEAGMNAICSSACGNPSPFKIQYATEPPEQSYLAIPASSWVDDFID 959

QY 960 WLTP-SSCCRLYISGPNKDKFCPSITVNSLNCILKNCMSITMGSVRPSVROFPHKYLWFLND 1018
Db 960 WLTP-SSCCRLYISGPNKDKFCPSITVNSLNCILKNCMSITMGSVRPSVROFPHKYLWFLND 1018

QY 1019 RNINIKCPKGLAAYSTSVNLTSQVYLASRFMAYHKPKNSQDYTEALRAARELANITA 1078
Db 1020 PNINICPKGLAAYSTSVNLTSQVYLASRFMAYHKPKNSQDYTEALRAARELANITA 1079

QY 1079 DLKRVGTPDPAPEVPPYITINVFYQYLTILPEGLPMLSLCLVPTFAVSCILGLDLRG 1138
Db 1080 DLKRVGTPDPAPEVPPYITINVFYQYLTILPEGLPMLSLCLVPTFAVSCILGLDLRG 1139

QY 1139 LNLLSIVMILVDTVGFMAWDISYNVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198
Db 1140 LNLLSIVMILVDTVGFMAWDISYNVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1199

QY 1199 ERAKATISMSGAVPAGVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLGLLHGLVFL 1258
Db 1200 ERAKATISMSGAVPAGVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLGLLHGLVFL 1259

QY 1259 PVLSYVGDVNPALALOKRAEBAVAAVWASCNHNSRVSTADNIYVNHSPESI-KG 1317
Db 1260 PVLSYVGDVNPALALOKRAEBAVAAVWASCNHNSRVSTADNIYVNHSPESI-KG 1317

QY 1318 AGAISNLFNPNNGROF 1332
Db 1317 ANAARSSLPKSDQKP 1331

RESULT 13
US-10-646-301A-2
; Sequence 2, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-646-301A-2

Query Match
Best Local Similarity 78.5%; Score 5421.5; DB 4; Length 1331;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSPPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 59

QY 61 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLASLSITKALLTRCPACSDN 119

QY 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVYAYEAFYQHSFAQSYDCSRVRVPA 179
Db 120 FVNLCHNTCSNQSLFINVTRVAQIGAGQLPAVYAYEAFYQHSFAQSYDCSRVRVPA 179

QY 180 RATLAVGTMCVYGSALCNAQWLNFGQDTGNGLAPLDTITFLLRPGQAVGSIQPLNEG 239
Db 180 RATLAVGTMCVYGSALCNAQWLNFGQDTGNGLAPLDTITFLLRPGQAVGSIQPLNEG 239

QY 240 VARNCSQDDVATSCDCAACPAIARPOALDSTFYLGMPGSLVLIILCSYFAVVT 299
Db 240 VARNCSQDDVATSCDCAACPAIARPOALDSTFYLGMPGSLVLIILCSYFAVVT 299

QY 300 ILLVGRVAPARDKSNVDPKGTSLSDKLSPTSTHLLGQPGQGWVWASWPLTILVLS 359
Db 300 ILLVGRVAPARDKSNVDPKGTSLSDKLSPTSTHLLGQPGQGWVWASWPLTILVLS 359

QY 360 VTPVVALAAGLVFTLITDTPVLSNAPSOARSEKAFHQHGGPFRFTNOVLTAPNRSS 419
Db 360 VTPVVALAAGLVFTLITDTPVLSNAPSOARSEKAFHQHGGPFRFTNOVLTAPNRSS 419

QY 420 KYRYSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSIQDICVAPLNPDN 479
Db 420 KYRYSLLGPKNFGSLDLDLLELLELQERLHQLVMSPEAQRNLSIQDICVAPLNPDN 479

QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDGTLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDGTLA 539

QY 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTFSLNYPAGDPRRLAQAKLWEAPLEE 599
Db 540 LSCMADYGAPVFPFLAIGYKGDYSEBALIMTFSLNYPAGDPRRLAQAKLWEAPLEE 599

QY 600 MRAPFORMAGMFQVFTFAERSLEDEINRTTADLPATSYIVIFLYISLALGYSYSSWR 659
Db 600 MRAPFORMAGMFQVFTFAERSLEDEINRTTADLPATSYIVIFLYISLALGYSYSSWR 659

QY 660 VVMSDKATLGLGGVAVLGVANMAAGFFSYLGRSLVLQVVPFLVLSVGADNIFIFVL 719
Db 660 VVMSDKATLGLGGVAVLGVANMAAGFFSYLGRSLVLQVVPFLVLSVGADNIFIFVL 719

QY 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFLGALTPMAVTRTFALTSLG 779
Db 720 EYQRLPRPGEPREVHIIGALGRVAPSMLLCSLSEAI CFFLGALTPMAVTRTFALTSLG 779

QY 780 VILDFLLQMSAFVALLSDSKQEARLDVCCVQPKQLPPQCGEGLLGFFQKAYAPF 839
Db 780 VILDFLLQMSAFVALLSDSKQEARLDVCCVQPKQLPPQCGEGLLGFFQKAYAPF 839
```

Qy	120	FVNHLCHNTCSNQSLSFINVTRVAQAGQALPAVVAYEAFYQHSFASQSDSCSRVRPA	179	Qy	1199	ERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLLITLGLHLGLVFL	1258
Db	120	FVSLHCHNTCSNQSLSFINVTRVVERGAGEPAVVAYEAFYQHSFASQSDSCSRVRPA	179	Db	1200	ERAKDAIVFMGSAFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLLITLGLHLGLVFL	1259
Qy	180	AATLAVGTWCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239	Qy	1259	PVILSYVGPDPNLPALAEQKRAEEAAVMVASCNPHSPSVSTADNIIYVNHSPGSI-KG	1317
Db	180	AASLAVGSMCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239	Db	1260	PVILSYVGPDPNLPALAEQKRAEEAAVMVASCNPHSPSVSTADNIIYVNHSPGSI-KG	1316
Qy	240	VARNCSQGDVATCSQDCAASCPAIAARPOALDSTFYLQMPGSLVLIILCSVFAVVT	299	Qy	1318	AGALSNFLPNNGROF	1332
Db	240	IAPCNESQGDVATCSQDCAASCPVIPPPEALRPSFYNGRMPGWLALIIIFTAVFVLLS	299	Db	1317	ANARSSSLPKSDQKF	1331
Qy	300	ILLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLS	359				
Db	300	AVLVRVVRNKNKAEGPQAPKLPKHKLSPHTILGRFFQNGMTRVASWPLTILVLS	359				
Qy	360	VIPVVALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419				
Db	360	FIWVIALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419				
Qy	420	YRYSILLGPKNFSGLDLDLLELELELELELELELELELELELELELELELELELELE	479				
Db	420	YKIDSILLGPKNFSGLDLDLLELELELELELELELELELELELELELELELELELELE	479				
Qy	480	TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTLA	539				
Db	480	TSLYDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTLA	539				
Qy	540	LSCMADYGAPVFPFLAIGYKGDYSEBALITWFSLNYPAGDPRLAQAKLWEAFLEE	599				
Db	540	LSCMADYGAPVFPFLAIGYKGDYSEBALITWFSLNYPAGDPRLAQAKLWEAFLEE	599				
Qy	600	MEAFORMAQMGVQFTABRSLEDEINRTIADLPFATSYIVIFLYISLALGSYSMSR	659				
Db	600	MESFORNTSKQVAFSAERSLEDEINRTIQDLPVFAVSIIYVIFLYISLALGSYSRCSR	659				
Qy	660	VMDSKATILGCGVAVVLGAVMAAMGFYSYLGIRSLVILQVVPFLVSVGADNIFIFVL	719				
Db	660	VAVESKATILGCGVAVVLGAVMAAMGFYSYLGIRSLVILQVVPFLVSVGADNIFIFVL	719				
Qy	720	EYQRLPRRGEPREVHIGALGEVAPSMLLCSLEAICFELGALTPMPAVRTFALTSGLA	779				
Db	720	EYQRLPRRGEPREVHIGALGEVAPSMLLCSLEAICFELGALTPMPAVRTFALTSGLA	779				
Qy	780	VILDFFLQMSAFVALLSLSKQESRLDVCCKVQPELPPQCGEGLLGFFQKAYAPF	839				
Db	780	IILDFFLQMTAFVALLSLSKQESRDPVLCFSTRKLPKPPKEGELLRLFRKIYAPF	839				
Qy	840	LLHWITRGVVLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYFVVGAP	899				
Db	840	LLHRTIRPVVMLFLTLFGANLYLMCNINVLGDLALPKDSVLLDYFLFLNRYLEVGP	899				
Qy	900	VYFVTILGYNFSSEAGNACSSAGCNFSFTOKIOYATEFPEOSVLAIPASSWDDFID	959				
Db	900	VYFVTISGYNFSSEAGNATCSAGCKSFSLTKIOYASEFPQSVLAIPASSWDDFID	959				
Qy	960	WLTP-SSCCLRYTSGPNKDKFCSTVNSLNCNLCMSITWGSVRPSVEQFHKYLPWFLND	1018				
Db	960	WLTPSSCCCLRYTSGPNKDKFCSTVNSLNCNLCMSITWGSVRPSVEQFHKYLPWFLND	1019				
Qy	1019	RPNKCPKGLAAYSTSVNLTSDGQVLASRFMAHYHKLKNSQDYTEALRAARELANITA	1078				
Db	1020	PPNIRCPKGLAAYSTSVNLTSDGQVLASRFMAHYHKLKNSQDYTEALRAARELANITA	1079				
Qy	1079	DLKRVCTDPAFEVFPPTIIVNFVQVLTTLPGELFMLSCLVPTPAVSCLLGLDLRS	1138				
Db	1080	DLKRVCTDPAFEVFPPTIIVNFVQVLTTLPGELFMLSCLVPTPAVSCLLGLDLRS	1139				
Qy	1139	LLNLLSVMLIVDTVGFMAWDISYNVAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWL	1198				
Db	1140	LLNLLSVMLIVDTVGFMAWDISYNVAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWL	1199				

RESULT 14

US-10-736-769-2
; Sequence 2, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-736-769-2

Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

Qy	1	MAEAGLRGWLWALLRLAQSEPYTHIQPGVCAYDECGKPELSGLMTLSNVSCLSN	60	Qy	1	MAEAGLRGWLWALLRLAQSEPYTHIQPGVCAYDECGKPELSGLMTLSNVSCLSN	60
Db	1	MAAAML-GWLLWALLSAAQGEIYTPKEAGVCTFYECGKPELSGLMTLSNVSCLSN	59	Db	1	MAAAML-GWLLWALLSAAQGEIYTPKEAGVCTFYECGKPELSGLMTLSNVSCLSN	59
Qy	61	TPARKITGDHLLTLOKICPRLYTGNTQ-ACCSAKQLVSEASLSITKALLTRCPACSDN	119	Qy	61	TPARKITGDHLLTLOKICPRLYTGNTQ-ACCSAKQLVSEASLSITKALLTRCPACSDN	119
Db	60	TPARVHTGEHLALORICPRLYNGNTTTPACSTQKLSLESMSITKALLTRCPACSDN	119	Db	60	TPARVHTGEHLALORICPRLYNGNTTTPACSTQKLSLESMSITKALLTRCPACSDN	119
Qy	120	FVNLHCHNTCSNQSLSFINVTRVAQAGQALPAVVAYEAFYQHSFASQSDSCSRVRPA	179	Qy	120	FVNLHCHNTCSNQSLSFINVTRVAQAGQALPAVVAYEAFYQHSFASQSDSCSRVRPA	179
Db	120	FVSLHCHNTCSNQSLSFINVTRVVERGAGEPAVVAYEAFYQHSFASQSDSCSRVRPA	179	Db	120	FVSLHCHNTCSNQSLSFINVTRVVERGAGEPAVVAYEAFYQHSFASQSDSCSRVRPA	179
Qy	180	AATLAVGTWCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239	Qy	180	AATLAVGTWCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239
Db	180	AASLAVGSMCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239	Db	180	AASLAVGSMCGVYGSALCNAORWLNFGQDTGNGLAPLDITFHLLERQOAVGSGIQPLNEG	239
Qy	240	VARNCSQGDVATCSQDCAASCPAIAARPOALDSTFYLQMPGSLVLIILCSVFAVVT	299	Qy	240	VARNCSQGDVATCSQDCAASCPAIAARPOALDSTFYLQMPGSLVLIILCSVFAVVT	299
Db	240	IAPCNESQGDVATCSQDCAASCPVIPPPEALRPSFYNGRMPGWLALIIIFTAVFVLLS	299	Db	240	IAPCNESQGDVATCSQDCAASCPVIPPPEALRPSFYNGRMPGWLALIIIFTAVFVLLS	299
Qy	300	ILLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLS	359	Qy	300	ILLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLS	359
Db	300	AVLVRVVRNKNKAEGPQAPKLPKHKLSPHTILGRFFQNGMTRVASWPLTILVLS	359	Db	300	AVLVRVVRNKNKAEGPQAPKLPKHKLSPHTILGRFFQNGMTRVASWPLTILVLS	359
Qy	360	VIPVVALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419	Qy	360	VIPVVALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419
Db	360	FIWVIALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419	Db	360	FIWVIALAAGLVFTLTDTPVELWAPNSQARSEKAFHDQHGPFRTNOVILTAPNRSS	419

Db	783	DFLLQMTAFVALLSLDSKQEARSPVWCFSRNLPPPKQKGBLLLCFFRKIYTPFLLH	842
Qy	843	WITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYEVPVGF	902
Db	843	RTIRPVLLFLVLFGANLYLNCNISVGLDQDUALPKDSYLLDYFLNRYLEVGPVGF	902
Qy	903	VTTLGVNFSBAGNNAICSSAGCNFNSFTQKIQYATEFFEQSYLAIPASSWVDDFDMLT	962
Db	903	DTTSGYNFSTEAGNNAICSSAGCESFSLQKIYASEFPNQSVALAASSWVDDFDMLT	962
Qy	963	P-SSCCRLXYISGPNKDKPCSTVNSNCLNCKNMSITWGSVRPSVEQFHXYLPFLNDRPN	1021
Db	963	PSSSCCRIYTRGPHKDFCFSTDTSFNCLNCKNMRITLGPVPTTEQFHXYLPFLNDRPN	1022
Qy	1022	IKCPKGGLAAYSTSVNLTSGQVLSRPFMAHYHKLXNSQDYTEALRAARELANITADLR	1081
Db	1023	IRCPKGGLAAYRTSVNLSDDQIIASQFMAYHKLPLRNSQDFTALRASALLAANITAEUR	1082
Qy	1082	KVPGTDPAFEVPPYTTINVFQYVLTILPGLFWMLSLCLVPTFAVSCLLGLDLRSGLN	1141
Db	1083	KVPGTDNFNEVPPYTTISNVFYQQYVLTVPGLFALCFVPTFVVCYLLGLDLRSGLN	1142
Qy	1142	LLSYIMILVDVTFGPMALWDISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLER	1201
Db	1143	LLSYIMILVDITGLMAYWNGISYNAVSLINLVAVGMSVEFVSHITRSFAVSTKPTLER	1202
Qy	1202	KEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIFFPRNLNLTLLGLLHGLVLPV	1261
Db	1203	KDATIFMGSAVFAGVAMTNPPGILILGFAQAQLIQIEFFRNLNLTLLGLLHGLVLPV	1262
Qy	1262	LSYVGPVNPALALEQKRAEBAVAVMVASCPNHPRSVSTADNIIYNHSPGSG- IKGAGA	1320
Db	1263	LSYLGPDVNQALVLEEKATEA-AMVSEPSQCQYFPFADANTSIDVYNGFNPFFIPEINA	1321
Qy	1321	ISNFLPNNGRQF	1332
Db	1322	ASSSLPKSDQKF	1333

Search completed: April 7, 2006, 19:37:11
Job time : 194 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:34:15 ; Search time 27 Seconds
(without alignments)
1538.804 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNGRQF 1332

Scoring table: BLOSUM62

Searched: 184161 segs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

.Listing first 45 summaries

- Database : Published Applications AA_New:*
1: /SIDSS/ptodata/1/pubppa/US08_NEW_PUB.pep:*
2: /SIDSS/ptodata/1/pubppa/US06_NEW_PUB.pep:*
3: /SIDSS/ptodata/1/pubppa/US07_NEW_PUB.pep:*
4: /SIDSS/ptodata/1/pubppa/PCT_NEW_PUB.pep:*
5: /SIDSS/ptodata/1/pubppa/US05_NEW_PUB.pep:*
6: /SIDSS/ptodata/1/pubppa/US10_NEW_PUB.pep:*
7: /SIDSS/ptodata/1/pubppa/US11_NEW_PUB.pep:*
8: /SIDSS/ptodata/1/pubppa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6896	99.8	1332	7	US-11-242-459-9 Sequence 9, Appli
2	2402.5	34.8	1278	6	US-10-995-561-952 Sequence 952, App
3	942	13.6	194	7	US-11-242-459-8 Sequence 8, Appli
4	195.5	2.8	1087	7	US-11-242-459-17 Sequence 17, Appl
5	195.5	2.8	1392	7	US-11-242-459-40 Sequence 40, Appl
6	191	2.8	127	7	US-11-096-568A-6778 Sequence 6778, Ap
7	173	2.5	85	7	US-11-096-568A-6779 Sequence 6779, Ap
8	143	2.1	445	7	US-11-242-459-35 Sequence 35, Appl
9	143	2.1	456	7	US-11-242-459-34 Sequence 34, Appl
10	141.5	2.0	1048	6	US-10-392-234A-18 Sequence 18, Appl
11	140.5	2.0	1046	6	US-10-392-234A-16 Sequence 16, Appl
12	134	1.9	1067	6	US-10-467-657-2548 Sequence 2648, Ap
13	127	1.8	1048	6	US-10-392-234A-14 Sequence 14, Appl
14	126.5	1.8	200	7	US-11-242-459-16 Sequence 16, Appl
15	126	1.8	801	6	US-10-793-626-2020 Sequence 2020, Ap
16	121.5	1.8	1043	6	US-10-392-234A-34 Sequence 34, Appl
17	120.5	1.7	2426	7	US-11-203-806A-11 Sequence 11, Appl
18	119	1.7	1048	6	US-10-392-234A-20 Sequence 20, Appl
19	117.5	1.7	1049	6	US-10-392-234A-12 Sequence 12, Appl
20	115.5	1.7	914	7	US-11-072-512-2923 Sequence 2923, Ap
21	113	1.6	1034	6	US-10-392-234A-30 Sequence 30, Appl
22	112.5	1.6	697	7	US-11-082-389-362 Sequence 362, App
23	109.5	1.6	758	7	US-11-096-568A-30412 Sequence 30412, A
24	109.5	1.6	772	7	US-11-096-568A-30411 Sequence 30411, A
25	109.5	1.6	827	7	US-11-096-568A-30410 Sequence 30410, A

ALIGNMENTS

RESULT 1

US-11-242-459-9
; Sequence 9, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9

Query Match	99.8%	Score	6896	DB	7	Length	1332
Best Local Similarity	99.8%	Pred. No.	0				
Matches	1330	Conservative	0	Mismatches	2	Indels	0
Gaps	0						
Qy	1	MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60				
Db	1	MAEAGLRGWLWALLRLAQSEPTTTHQPGYCAFYDECGKNPGLSGSLMTLSNVSCLSN	60				
Qy	61	TPARKITGDHLILLOKICPRLYTGTNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF	120				
Db	61	TPARKITGDHLILLOKICPRLYTGTNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF	120				
Qy	121	VNLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAEYAFYQHSFAEQSVDCSRVRVPA	180				
Db	121	VNLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAEYAFYQHSFAEQSVDCSRVRVPA	180				
Qy	181	ATLAVGTMCYGVYGSALCNQAWLNFGQDTGNGLAFLDITFHLLRPGQAVGSIQPLNEGV	240				
Db	181	ATLAVGTMCYGVYGSALCNQAWLNFGQDTGNGLAFLDITFHLLRPGQAVGSIQPLNEGV	240				
Qy	241	ARCNSQGDVATCSQDCAASCAPAIARPOLDSTFYLGQMPGSLVLIILCSVFAVVTI	300				
Db	241	ARCNSQGDVATCSQDCAASCAPAIARPOLDSTFYLGQMPGSLVLIILCSVFAVVTI	300				

301 LLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
301 LLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV 360
361 IPVVALAAGLVFTLTTDPVBLWSAPNSQARSEKAFHDQHPFPRTNQVLTAPNRSY 420
361 IPVVALAAGLVFTLTTDPVBLWSAPNSQARSEKAFHDQHPFPRTNQVLTAPNRSY 420
421 RYDSLGLGKPFSGTLDLLELELELERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
421 RYDSLGLGKPFSGTLDLLELELELERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGATLAL 540
481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGATLAL 540
541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
601 RAQORMAGMFOVTTFAERSLDEINRTTAEDLPFATSYIVIFLYISIALGSSYSWSRV 660
601 RAQORMAGMFOVTTFAERSLDEINRTTAEDLPFATSYIVIFLYISIALGSSYSWSRV 660
661 MYDSEKATLGLGGVAVVLGAVMAAGFFSYLGRSSLSLVLQVVPFLVSGADNIFIFVLE 720
661 MYDSEKATLGLGGVAVVLGAVMAAGFFSYLGRSSLSLVLQVVPFLVSGADNIFIFVLE 720
721 YQRLPRRPEPEVHIGRALGVAPSMLLCSLSEACPELGAITMPAVRTEALTSLAV 780
721 YQRLPRRPEPEVHIGRALGVAPSMLLCSLSEACPELGAITMPAVRTEALTSLAV 780
781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCKVQQLPPQGGEGLLGFPKAYAPPL 840
781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCKVQQLPPQGGEGLLGFPKAYAPPL 840
841 LHMTRGVVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFEVGA 900
841 LHMTRGVVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFEVGA 900
901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSTOKIYATFPEOSYLAI PASSWDDFIDW 960
901 YFVTTLGYNFSSEAGNNAICSSAGCNPFSTOKIYATFPEOSYLAI PASSWDDFIDW 960
961 LTPSSCCRLYISGPNKDKFCPTSVNSLNCNKMCSITMGSVRPSVBFHXYLPWFND 1020
961 LTPSSCCRLYISGPNKDKFCPTSVNSLNCNKMCSITMGSVRPSVBFHXYLPWFND 1020
1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKLNSQDYTEALRAARELANITADL 1080
1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKLNSQDYTEALRAARELANITADL 1080
1081 RKVPGTDPAEVFPYITNTVYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
1081 RKVPGTDPAEVFPYITNTVYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
1141 NLLSIVMLVDTVGFMAWLDISYNAVSLNLVSAVGSVEFVSHITRSPAISTKPTWLER 1200
1141 NLLSIVMLVDTVGFMAWLDISYNAVSLNLVSAVGSVEFVSHITRSPAISTKPTWLER 1200
1201 AKENATISMGSAVAGVANTMLPGLVLGLAKAQIQLIFFRLNLLITLGLLHGLVFLPV 1260
1201 AKENATISMGSAVAGVANTMLPGLVLGLAKAQIQLIFFRLNLLITLGLLHGLVFLPV 1260
1261 ILSVVGDPVNPALAEOKRAEEAAVAVMVASCPNHPRSVSTADNIYVNHSEFSGIKAGA 1320
1261 ILSVVGDPVNPALAEOKRAEEAAVAVMVASCPNHPRSVSTADNIYVNHSEFSGIKAGA 1320
1321 ISNFLPNNRGQF 1332
1321 ISNFLPNNRGQF 1332

RESULT 2
US-10-995-561-952
; Sequence 952, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 952
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-952

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.6e-181;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMILLMALLRLAQSEPVTTIHQPGYCAFYDECG-----KNPELSGLMTLSNVVCLS 59
DB 4 RGLALGILLLLLLCPAQVFSQ-----SCWYGECCGLAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLILLOKICPRLYTGPNTPQACCSAKQLVSLBASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQLQTLKDNLQLPLQFSLRCPSPFN 103
QY 120 FVNLHCHTCSFNQSLFINVTR-----VAQLGAGQLPAVAVAEAFVQHSFASQSDSCSRV 175
DB 104 LNLFLCELTCSPROSFQFNVLTATEDYVDPVTNQTKTNVKELYQYVQGSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCQVYGSALCAQRWLNFGDGTGNGLAPLDIT-----PHLEPQQA 228
DB 164 EAPSSNDKALGLLCKGDADA--CNATNWEYFMKNDGQAPFTITVPVSDFPVH----- 215
QY 229 VGSQIQLNEGVARCNESQDDVATCSQDCQCAASCPAARPO-----ALDSTFYLG 279
DB 216 ---GMEPNWNAKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPPTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSVFAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
DB 273 WITYMAFLLVFFGAFVAVCWYKRYFVSEYTPIDNSIAFSV--NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFFQGGTGWASWPLTILVLSVIVVVALAAGLVFTLTTDPVBLWSAP 386
DB 325 DPVSAAFEGCLRLFTRWGSCFVRNPGCVIPFSLVFTACSSGLVFRVVTNTPVDLWSAP 384
QY 387 NSQARSEKAFHDQHPFPRTNQVLTAPNRSYRSDLSLLGPK--NFSGLDLDLLELL 445
DB 385 SSOARLEKEYFDQHPGFPFTEQLIRAPLTDKHIYQYPYSGADVFGPPLDIQILHQL 444
QY 446 ELQERLHLQVMSPEAQRNLSLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTIQLDCLAPLSPYNT---NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDHFLYCANAPLTFKDGATLALSCMADYGAPVFPFLAIGGYKGDYS 565
DB 500 GDDFF---VYADVHTFLYCVRAPASLNDTSLDHPCLGTGPGVFPVFWLVLGGYDQNYN 556
QY 566 EBAALIMTSLNNYPAGDPRLAQAKLWEEAFLEEMBAFORWAGMFQVFTFAERSLDEI 625
DB 557 NATALVITFPVNNYNDTEKLRQAQAWKEKFINVKNYKN---PNLTIESTAERSLEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYISIALGSSYSWSRVVMDSKATLGLGGVAVVLGAVMAAG 685
DB 614 NRESDSDVTVVISVAIMFLYISLALGHIKSCRLLVDSKVLGIAGILIVLSSVACSLG 673

QY 686 FFSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRPGEPRVHIGRALGRVAP 745
Db 674 VFSYGLPLTVLIVIEIPFLVLAAGVDNIFILVQAYQDERLQGETDQDLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSSFSFETVAFGLGALSMVPAVHTFSLPAGLAVFIDFLLQITCFVSLGLDIKRQEKV 793
QY 806 RLDDVCCVQQLPPPGQ-GEGLLGFQKAVAPFLLHWTIRGVVLLFLFALFGVSLYSM 864
Db 794 RLDFCCVRGAEDGTSVQASESCLFPPKNSYSPILLKDWMPVIAIFVGVLSIAVL 853
QY 865 CHISVGLDQELAPKDSYLLDFLFLNRYFEVQAPVYFTTILGYNPSSAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDSDYMDYFKSISQYLHAGPPVYFVLEBHDYTSKGGQNMVGGMG 913
QY 925 CNFNSFTQIQVATSPPEQSYLAIPASSWVDPIFWLTP-SGCCRLYISGPNKDKPCPT 983
Db 914 CNNDLSVQIFNAQOLDNRYTRIGPAPSSWIDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT-MGSVRSVPEQPHKYLPWFLNDRPNKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTEGKQRPQGGDFMFLPFLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQCVLASRPMAYHKLKNSQDTEALRAARELAANITADLRKVPGTDPAFEVPPYITNV 1100
Db 1030 GTRVGATYEMTYHTVLTQTSADPIDALKKARLIASNT-ETMGINGS--AYRVEPYSVFY 1086
QY 1101 FVEQVLTILPEGLFMLSLCLVTPFAVSCLLGLDLSRGLNLLSVMIILVDTVGFMAWD 1160
Db 1087 FVEQVLTIIDDFINLGVSLGAIFLVTWVLLGCELWSAVIMCATAMVLVNMFGVWLMG 1146
QY 1161 ISYNAVSLINLAVSGMSVEFVSHITRSPAIKPTWLERAKEATISMSAVPAGVAMTN 1220
Db 1147 ISLNAVSLVNLVWMSGISVEPCSHITRAFTVSMKGSRVERAEALAHMGSVFSGITLK 1206
QY 1221 LRGILVLGLAKQLIQIFPRNLNLLITLLGLHLVFLPVLISVYGVDPVNP 1272
Db 1207 FGGIVVLAPAKSQIFQIFRMYLAMVLGATHGLIFLPLVLSYIGPSYNKA 1258

RESULT 3

US-11-242-459-8
; Sequence 8, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 8
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-8

Query Match 13.6%; Score 942; DB 7; Length 194;

Best Local Similarity 99.5%; Pred. No. 5.1e-67;

Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 615 FTAESLEDEINRTTAEDLPATATSYIVIFLYISALGYSWSRWVDSKATLGLGVA 674

Db 1 FMAESLEDEINRTTAEDLPATATSYIVIFLYISALGYSWSRWVDSKATLGLGVA 60

QY 675 VVLGAVMAAMGPFSSYLGRSSLVILQVVPFLVSVGADNIFIFVLEYQRLPRPGEPRV 734
Db 61 VVLGAVMAAMGPFSSYLGRSSLVILQVVPFLVSVGADNIFIFVLEYQRLPRPGEPRV 120
QY 735 HIGRALGRVAPSMMLCSLSEATCFPLGALTTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 794
Db 121 HIGRALGRVAPSMMLCSLSEATCFPLGALTTPMPAVRTFALTSLGLAVILDFLLQMSAFVAL 180
QY 795 LSLDSKROEASRLD 808
Db 181 LSLDSKROEASRLD 194

RESULT 4

US-11-242-459-17
; Sequence 17, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; PRIOR FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-17

Query Match 2.8%; Score 195.5; DB 7; Length 1087;

Best Local Similarity 19.5%; Pred. No. 6.1e-07;

Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEENRAFQRMAGMFQVFTTAERSLED----- 623

Db 100 LPNTYSVDDR-----WEE-----QRAKFSQSVTVYVAVLAKQSTSKVQVLYGGTD 144

QY 624 ----EINRTTARD--LPITATSYIVIFLYISALGYSWSRWVDSKATLGLGG-VAVV 676

Db 145 LFDYEVRETRFNNMMLAPISSSCIAALVYILTSCSVFLSFFGI-----ASIGLSCLVALP 199

QY 677 LGAVMAAMGPFSSYLGRSSLVILQVVPFLVSVGADNIFIFVLEYQRLPR-RPGEPRVH 735

Db 200 LYHVVFGEI---QYLGILNG-----VAAPVIGIGVDDVVFVINTYRQATHLEDPOLRMTH 251

QY 736 IGRALGRVAPSMMLCSLSEATCFPLGALTTPMPAVRTFALTSLGLAVI---LDLFLQMSAFV 792

Db 252 TVQTAGK---ATFTSLTTAAAYANVFSQIPAVHDFGLPWSLIVSCCWLAIVTMPAL 308

QY 793 ALLSLDSKROEASRLDVC--CCVKPQELPPPGQ----- 823

Db 309 GLWSLYLAPLESSCQTSCHQNSRKTSLHFPDGVFAAPEQVGGSPAQGPPIPYLDDDDIPL 368

QY 824 -----GEGILLGFQKAYAP-----FLJHMI-----TRGVVL 850

Db 369 EVEEPVSELEIGDVSLSVSPGLOPASNTGSRGHLIVOLQELHHHWMWSAVKSRWIV 428

QY 851 LFLTALFGVSL-----YSM-----CHISVGLDOE- 874

Db 429 GLFVSILILSVFASRLRPASRAPLLFRPDNIQVLLDLKYNLSAEGISCITCSGLPQEK 488

QY 875 -----LALPK-----DSYLLDY--FLPLNRYFEVQAPVYFVTTLGVNF 910

Db	489	PHSLQNNIRTSLEKRRGSGVFWASREPEATLQDPFGTYI:SKVKSQGHPA--VYRLSLNA	546
Qy	911	SSEAGMAICSSAG--CNMF-----SFTOKIQVATEFFPQSLALPASPSSVDDFDLWL	961
Db	547	SLPAPMQAVSGDGEVPSFQVYRAPFGNFTKLTACHSTVGLLQAASPSRKM-----L	600
Qy	962	TPSSC-----CRLVIS--GP	974
Db	601	TTLACDAKRGWKDFSFYVATKEQHTKRLYFAQSHKPPPHGRVCMAPPGLSSPDGP	660
Qy	975	NKDFCPSTVNSLNCNCSITWG-----SVRPSVEQPHKLPWF-----LN-	1017
Db	661	TKGFF---FVPSEKVPKARLSATFGFPCVNTGCGKPAVRPLVDTGAMVFFVGIIGVNR	717
Qy	1018	---DRPNKCPKGLAAYSTSVNLTS-----GOVL	1045
Db	718	TRQVDNHHVIGDP--GSVYDSSFDLFEKIGHLCHLCKAIAANSELVKPGGAQLPSGYSI	775
Qy	1046	ASRFMAYHKPLKNSOD-----YTEALRAARELANITADLRKVPGTDPAPFVP-P	1094
Db	776	SSFLQMLHPECKELPEPNLLPGQLSHGAVGVREGVQWISMAFESTTYKGS-SFQYSD	834
Qy	1095	YITITNVFEQYLTILPEG-----LFW-----LSLCL--VP	1122
Db	835	YLRWESFLQQLQALPEGSVLRGFGTCEHWKQIFMEIVGVQSALCGLVLSLLICVAAVA	894
Qy	1123	TFVAVSLGLDLRSGLNLLSIVMLVDVTFGMALWDISYNAVSLINLVSAVGMVBVF	1182
Db	895	VFTTHILL-----LLPVLSILGIVCLV--TIMYWSGEM--GAVEAISLSILVGSVDYC	947
Qy	1183	SHITRSFAIS-----TKPTWLERAKEATISNGSAVFAVAGVAMTNLPGI-----	1224
Db	948	VHLVEGYLLAGENLPPHQAEDARTQW--RTLEAVRHVGVAIVSSALTITVATVPLPFC	1005
Qy	1225	LVLGLAKAQIQLIFFPRMLNLLTLGLLHGLFPLVILSVGPDVNPALALEQKRAE--E	1282
Db	1006	IIAPFAK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK	1053
Qy	1283	AVAAVMVA	1290
Db	1054	ALGAVLLA	1061
RESULT 5			
US-11-242-459-40			
; Sequence 40, Application US/11242459			
; Publication No. US20060035835A1			
; GENERAL INFORMATION:			
; APPLICANT: TANIYAMA, Yoshio			
; APPLICANT: KITA, Shunbun			
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof			
; FILE REFERENCE: 2703USOP			
; CURRENT APPLICATION NUMBER: US/11/242,459			
; CURRENT FILING DATE: 2005-10-03			
; PRIOR APPLICATION NUMBER: US/10/239,316			
; PRIOR FILING DATE: 2002-09-19			
; PRIOR APPLICATION NUMBER: PCT/JP01/02279			
; PRIOR FILING DATE: 2001-03-22			
; PRIOR APPLICATION NUMBER: JP2000-088595			
; PRIOR FILING DATE: 2000-03-24			
; NUMBER OF SEQ ID NOS: 59			
; SEQ ID NO 40			
; LENGTH: 1392			
; TYPE: PRT			
; ORGANISM: Human			
US-11-242-459-40			
Query Match 2.8%; Score 195.5; DB 7; Length 1392;			
Best Local Similarity 19.5%; Pred. No. 8.6e-07;			
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;			
Qy	576	LNNYPAGDPRLAQAKLWEEAFLEWRAPFORMAGNFQVTFTAERSLED-----	623

Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEENRAFQRRMAGMFQVTFTAERSLED----- 623
DB 100 LPNYSVDDR-----WEE-----QRAKFSQFVVTVVAMLAQKSTSKVQVLYGGTD 144

QY 624 ----EINTTAED--LPFATSYIVIFLYISALGSYSSWSRVMVDSKATLGLGG-VANV 676
DB 145 LFDVEVRTFNNDMLAPFISSCIAALVYILTSCVFLSPFGI-----ASIGLSCLVALF 199

QY 677 LGAYMAAGFFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLEYQRLPR-RPGEPREVH 735
DB 200 LYHVVFGI---QYLGILG-----VAAFVIVGIGVDDVFVINTYRQATHLEDQLRWIH 251

QY 736 IGRALGRVAPSMMLCSLSEACFFLGALTMPMPAVRTFALTSLGLAVI---LDFLQMSAFV 792
DB 252 TVQTAGK---ATFTSLTAAAYANVFSQIPAVHDFGLFSLIVSCCMLAVLVTMPAAL 308

QY 793 ALLSLDSKRQEASRLDVC---CCVKQELPPPG 822
DB 309 GLWSLYLAPLESSCQTSCHQNCRSKRTSLHFFG 340

RESULT 10
US-10-392-234A-18
; Sequence 18, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxeer, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-392-234A-18

Query Match 2.0%; Score 141.5; DB 6; Length 1048;
Best Local Similarity 18.1%; Pred. No. 0.011;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;

QY 624 EINTTAEDLPFATSYIVIFLYISALGSYSSWSRVMVDSKATL--GLGGVAVILG--A 679
DB 339 EVVKTLVE-----AIIIVFLWYLFQNF-----RATLPTIAVPVLLGTEA 381

QY 680 VMAAMGFFSYLGRSSILVQVFLVLSVG--ADNIFVLEYQRLPRRPGEPREHVIG 737
DB 382 VLAAGFG-----SINTLTMFG--MVLAIGLLVDDAIVVENVERVMTTEEGLPKREATR 432

QY 738 RALGR-----VAPSMILCSLSEACFFLGALTMPMPAVRTFALTSLGLAVILDFLLQMSAFV 792
DB 433 KSMGQIOGALVGIAMVLSAVFIPWAFPGSGTGAI--YRQFSITIVSMALSVLVALILTP 490

QY 793 ALLSLDSKRQEASRLDVCCKVQELPPPGQGLLIGFFQKAYAPFLHVI----- 844
DB 491 ALCA-----TMLKPVAKDGHGKKGFGWENRFLFDKSTHTHYTDSVGNILR 536

QY 845 TRGVLLFLALF-----QVLSYMSCHISVGLDQLALPKDSYLL 884
DB 537 STGRYLLLYIIIVGMAYLFVRLPSSFLPDEDQGVFL--TWQVPAGATQERTOKVLDEVT 595

QY 885 DYFL-----FLNR----- 892

DB 596 DYVLNKEKANVESVPAVNGFGFAGRQNTGIAFVSLKMDWDRPGEKKNKVEAITQRATAAF 655

QY 893 -----YFVGCAPVYFVTTLGYNFSSEAGNAICSSAGCNCNFSFTQ----- 932

DB 656 SOIKDAMVAFNLPAIVELGTATGDFPE-----LIDQAGLGEKUTQARNQLFGEVAK 708

QY 933 -----KIQVATEPEQSYLAIPAS-----SWVDDPFDWLTP 963

DB 709 YPDLLVGVRNGLEDTPQKIDIDQKQAQALGVISIDINTTLGAAMWGSVYVNDPID--R 765

QY 964 SSCRLYISGNPKDKFCFSTVNSLNCNLCMSITMGSVRSVQFHKYKLPFLNDRNIK 1023

DB 766 GRVKKVVVMSEAKYRMLPDDIND-----WYVRSQDQM 798

QY 1024 CPKGLAAYST-----SVNLTSQGVLASRFWAYHKLKNSQDYTEALRAARELANIT 1077

DB 799 VP---FSAFSSRWYSGSPRLRYNGLPSMEILGQAAPGKST---GEAMAMEELASKLP 852

QY 1078 ADLRKVPGTDPAPFVFPYITITNVFYEQLT-----ILPEGLFMLSCL----- 1120

DB 853 SGI-----CYDWTGMSYQERLSGNQAPALYALISLVVFLCLAALYESWSIPF 899

QY 1121 ---VPTFAVSCLLGLDLRSGLLNLISIVMILVDTVGMALWDISYNAVSLN----- 1170

DB 900 SVMLVPLGVIGALLAA--TFRGLTNDVYFQVGLLTIGLSA-----KNAILIYEFKDL 952

QY 1171 -----LVSAGVMSVEFVSHITRSPAISKTPTWLERAKEATISMGSAVPAVAVMT 1219

DB 953 MDKEGKGLVEAMLEAVRMRLPILMTSLAFMLGVP-----LVISSG-----AGSGAQ 1000

QY 1220 MLPGLVL--GLAKAQLIQIFP 1239

DB 1001 NAVGTGVLGVMVATVLAIFP 1021

RESULT 11
US-10-392-234A-16
; Sequence 16, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxeer, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-392-234A-16

Query Match 2.0%; Score 140.5; DB 6; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.013;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

QY 506 NOTLMGOT--SQVDWKDHFYCANAPLTPKDGTFALASCADYAGVAPFPFLAIGGYKGD 563

DB 231 NATIIGTKLQTAEPENILLKVN-----PDGSOVRLKDVADV-----LG---GDD 274

QY 564 YSEAEALIMTFLNNYPA-----GDPRLAQAKLWEE--AFLEEMRAFQRRMAGMFQ 612

DB 275 YS-----INAQFNGSPASGIAIKLATGANALDTAKAIRQTIANLEFPFQGMKVVPYD 328

QY 613 VFTPAERSLEDEINRTAEDLPFATSYIVIFLYISALGSYSSWSRVMVDSKATL--GL 670

Db 329 TTPVVSASTH-EVVKTLGE-----AIIIVFLVMYLFLONF-----RATLIPTI 370
QY 671 GGVAVVLG--AVMAAMGFSSYLGRSSVLQVVPFLVLSVG--ADNIFIFVLEYQRLPR 726
Db 371 AVPVVLGTFGVLAAGF-----SINTLTWFG-MVLAIGLLVDDAIVVVVENVERVA 421
QY 727 RQG-EPREV--HIGBALGR-VAPSMLLCSLSEAICFFLGALTPMPAVTTPALTSLGLAVI 781
Db 422 BEGLSPREARKSMQIOGALVGIAMVLSAVFLPMAFPFGSGTGI--YRQFSITIVSAMA 479
QY 782 LDFLLQMSAFVALLSLDSKRQASRLDVCC--VKPQLPPPGQEGGLLGFQOKAYAP 838
Db 480 L-----SVIVALI-----LTPALCATMLKPIEKDGHGKGFPGFNKMFUS 522
QY 839 FLLHMITRGV-----LILLFLALFGVSLYSMCHISVGL-----DQ-----ELALPKD 880
Db 523 -TTHGVERGVASILKHRAFPVLLIYVIVVAGMIWMFTRIPTAFLPDEDQGVLAQVQTPPG 581
QY 881 S-----YLLDYFLFLNRYFVGAPVFPVTTGLYNFSSE----- 913
Db 582 SSAERTQVVVDSMRLEYLLEK-----ESSVSVVFTVTGTFAGRGQSSGMAFIMLXP 633
QY 914 ----AGWNAI---CSSACGNFSTQKIQYATEPEQSYLAIPASSWVDDFID----- 959
Db 634 WEERPGGENSVFELAKRAQMHFFSFKDAWVFA--FAPPSVLBELGNATGDFLQDOAGVG 691
QY 960 ----WLTSSCCRLYISGPNKDKFCPSTVN-----SLNCLNCKMSI 996
Db 692 HEVLQARNKFLMLAAQNPAQORVRPNMGMSDEPOYKLEIDDEKASALGVSLADINSTVSI 751
QY 997 TWGSV-----RPSVEQPHKYLPHLNDPN-----TKCPKGGIA---AYST 1034
Db 752 AWGSSVYVNDFIDRGVRV-----YLOGRPDARWNPDDLKSWYVRNDRKGMVFPNAPAT 805
QY 1035 ----SVNLTSDGQVLASRFWAYHKPLKNSODYTEALRAARELAANITADLRKVPCTDP 1088
Db 806 GKWEVGSPLERYNGVPAMEILGEAPGLSSGD---AMAAVEE---IVKQLPKGVG--- 855
QY 1089 APEVFPYTTITNVFEYOYIT-----ILPEGLFMLSIC-----LVPTFAYS 1127
Db 856 ----YSWTGYSYERLSSGQAPALYALSLLVFLCLAALYESWSIPPSVMLVPLGTI 909
QY 1128 CILLGLDLRSGLNLLSIVMILVDTVGFMAWDISYNVSLINL-----V 1172
Db 910 GALLATSMR-GLSNDVFFQVGLLTTIGLSA-----KNAILVEFAKELHEQCKGIVEAAI 963
QY 1173 SAVGMSVEFVSHITSFALSTPTWLERAKEATISMGSAVFACVAMTNLPGLVL-GLAK 1231
Db 964 EACRMRLRPVMTSLAFLIGVVP-----LAISTG-----AGSGSQHAIGTGVIGGMVT 1011
QY 1232 AOLIQIFPRL 1242
Db 1012 ATVLAIFWVPL 1022

RESULT 12

US-10-467-657-2648

; Sequence 2648, Application US/10467657

; Publication No. US20050260581A1

; GENERAL INFORMATION:

; APPLICANT: CHIRON Spa

; APPLICANT: FONTANA Maria Rita

; APPLICANT: PIZZA Mariagrazia

; APPLICANT: MASIGNANI Vega

; APPLICANT: MONACI Elisabetta

; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/467,657

; CURRENT FILING DATE: 2003-08-11

; PRIOR APPLICATION NUMBER: GB-0103424.8

; PRIOR FILING DATE: 2001-02-12

; NUMBER OF SEQ ID NOS: 9218

; SOFTWARE: Seqwin99, version 1.04

; SEQ ID NO 2648

; LENGTH: 1067

; TYPE: PRT

; ORGANISM: Neisseria gonorrhoeae

US-10-467-657-2648

Query Match

Best Local Similarity 19.4%; Score 134; DB 6; Length 1067;

Matches 185; Conservative 139; Mismatches 354; Indels 274; Gaps 44;

QY 452 RHLQW-SPEAR--NISLQDICYAPLNPDPNTSLYDCCINSILQYQNNRTLLLLLTANQT 508
Db 180 RAMRIVDPKCLQNTNLSPADVGS--LSAQNQISAGSIGSLPAV--RQGT---VATATV 233
QY 509 LMGQTSQVDKDHFLYCANAPLTFKQGTALALSCMADYCAPVFPPLAIGCYGKDYSEAE 568
Db 234 AQOQLGTAEFGNVILRANT-----DGSNIYLDKVAKVGLM-----EDSSST 277
QY 569 ALI-----MTFSLNNYPAGDPRLAQAQKLWEBAFLBEMRAFORRMAGMFQVTTPTABRSLE 622
Db 278 RLNGVNTTGMVAMLSN--SGN-AMATAKAVKERLAVLEKYFPQGMS--WKTPTDTSKFVE 332
QY 623 DEINRTAEDLPIPATSYIVIFLYISLALGSSYSSSRVWVDSKATLGLGVAVVLGAVMA 682
Db 333 ISIEKVIHTLIBAMVLVFMVYLFQNI--RYTLPTIIVVPISL---LGGFA----- 379
QY 683 AMGFPSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEBREHVHIGRALGR 742
Db 380 ---FISYGM--SINVLTPMAMILVIGIVDDAIVVVENVERIMAGEGLPPKEATKKAMQ 435
QY 743 VAPSM-----LCSLSEAICFFLGALTPMPAVTTPALTSLGLAVILDFLQMSAFVALLSL 797
Db 436 ISGAVIGITAVLISVFPVPLMSPGGAAGNI--YKQFALTWASSI-----AFSAFLA--- 484
QY 798 DSKRQASRLDVCCCKVQELPPPGQGGEL--LLGFFQKAYAPF-----LHWT 845
Db 485 ----TLTPALCATMLKTIPIKGHBEKKGFGFNKPKPDSWTHYGEGRVAKVLKTF 536
QY 846 RGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPVFPVT 905
Db 537 RMVYVIGLVVGV-----FLFMR-----LPTSFLPT 563
QY 906 LGYNFSBAGMNAICSSAGC-----NNFSPTQKIQYATEPPE--QSYLAIPASSWVDD--- 956
Db 564 EDQGFV---MVSVOLPAGATKERTDATALAQVTLAKSIPEIENIITVSGFSGSQNM 619
QY 957 ----FDW--LTPSSCCRLYIS-----GPNKDKFCPSTVN-----S 986
Db 620 AMGFALKDWNBERTASGSDAVAVAGKLTGMMGTLDKDFGISVWVPPPILLEGNGSLIN 679
QY 987 LNCLNCKMSITMGSVRPSVEQPHKYLPMF-----LNDPNKCPKGGIAAYSTSVN 1037
Db 680 LQDRNNTGHTALLAKRNELIQMRASGLFPDPTVTRAGGLEDSFQKIDINRAAAQAQGIS 739
QY 1038 LTSQGVLASRF-MAYHKPLKNSODYTEALRAARELAANITADLRK-----VP--- 1084
Db 740 FADIRITALASALSSSYVDFPNQGRQVRWVQADEADARMQPADILNLTPNKSQVAVPLS 799
QY 1085 ----GTD-----PAFEVFPYTTITNVFEYOYI----- 1106
Db 800 TIATVSWENGTEQSVRFNGYPSMKLSASPATGVTGSTQMAAAVQKMWVDELGGYSPEWGO 859
QY 1107 ----TILPEGLFMLSICLV-----PTFAVSCLLAGL-----DLR 1136
Db 860 SREEAKGSGQTLILYGLAVAAVFLVLAALYSWSIPLAVILVPLGLTGAAAGVTGRNLF 919
QY 1137 SGLL-NLLSIVMILVDTVGFMAWDIS--YNVSLINLV-----SAVMSVE-----F 1181
Db 920 EGLGVSFSPANDIYFQVGFVTVMLGSAKNAILIEPAKDLQAQCKSAVEAALEAARLP 979
QY 1182 VSHITRSPA--ISTKPTWLERAKEATI--SMGSAVFAGVAMTNLPGLVLGL 1229
Db 980 RPIIMTSPAFILGVVPLVIAAGASSASQRAIGTTVFWGMLVGTLLSVFLVPL 1031

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:22:20 ; Search time 569 Seconds
(without alignments)
3235.086 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main:

- 1: /cgm2_6/ptodata/1/paa/PCTUS_COMB.pep:**
- 2: /cgm2_6/ptodata/1/paa/US066_COMB.pep:**
- 3: /cgm2_6/ptodata/1/paa/US073_COMB.pep:**
- 4: /cgm2_6/ptodata/1/paa/US074_COMB.pep:**
- 5: /cgm2_6/ptodata/1/paa/US075_COMB.pep:**
- 6: /cgm2_6/ptodata/1/paa/US076_COMB.pep:**
- 7: /cgm2_6/ptodata/1/paa/US077_COMB.pep:**
- 8: /cgm2_6/ptodata/1/paa/US078_COMB.pep:**
- 9: /cgm2_6/ptodata/1/paa/US079_COMB.pep:**
- 10: /cgm2_6/ptodata/1/paa/US080_COMB.pep:**
- 11: /cgm2_6/ptodata/1/paa/US081_COMB.pep:**
- 12: /cgm2_6/ptodata/1/paa/US082_COMB.pep:**
- 13: /cgm2_6/ptodata/1/paa/US083_COMB.pep:**
- 14: /cgm2_6/ptodata/1/paa/US084_COMB.pep:**
- 15: /cgm2_6/ptodata/1/paa/US085_COMB.pep:**
- 16: /cgm2_6/ptodata/1/paa/US086_COMB.pep:**
- 17: /cgm2_6/ptodata/1/paa/US087_COMB.pep:**
- 18: /cgm2_6/ptodata/1/paa/US088_COMB.pep:**
- 19: /cgm2_6/ptodata/1/paa/US089_COMB.pep:**
- 20: /cgm2_6/ptodata/1/paa/US090_COMB.pep:**
- 21: /cgm2_6/ptodata/1/paa/US091_COMB.pep:**
- 22: /cgm2_6/ptodata/1/paa/US092_COMB.pep:**
- 23: /cgm2_6/ptodata/1/paa/US093_COMB.pep:**
- 24: /cgm2_6/ptodata/1/paa/US094_COMB.pep:**
- 25: /cgm2_6/ptodata/1/paa/US095_COMB.pep:**
- 26: /cgm2_6/ptodata/1/paa/US096_COMB.pep:**
- 27: /cgm2_6/ptodata/1/paa/US097_COMB.pep:**
- 28: /cgm2_6/ptodata/1/paa/US098_COMB.pep:**
- 29: /cgm2_6/ptodata/1/paa/US099_COMB.pep:**
- 30: /cgm2_6/ptodata/1/paa/US100_COMB.pep:**
- 31: /cgm2_6/ptodata/1/paa/US101_COMB.pep:**
- 32: /cgm2_6/ptodata/1/paa/US102_COMB.pep:**
- 33: /cgm2_6/ptodata/1/paa/US103_COMB.pep:**
- 34: /cgm2_6/ptodata/1/paa/US104_COMB.pep:**
- 35: /cgm2_6/ptodata/1/paa/US105_COMB.pep:**
- 36: /cgm2_6/ptodata/1/paa/US106_COMB.pep:**
- 37: /cgm2_6/ptodata/1/paa/US107_COMB.pep:**
- 38: /cgm2_6/ptodata/1/paa/US108_COMB.pep:**
- 39: /cgm2_6/ptodata/1/paa/US109_COMB.pep:**
- 40: /cgm2_6/ptodata/1/paa/US110_COMB.pep:**
- 41: /cgm2_6/ptodata/1/paa/US111_COMB.pep:**
- 42: /cgm2_6/ptodata/1/paa/US112_COMB.pep:**
- 43: /cgm2_6/ptodata/1/paa/US114_COMB.pep:**

- 44: /cgm2_6/ptodata/1/paa/US600_COMB.pep:**
- 45: /cgm2_6/ptodata/1/paa/US601_COMB.pep:**
- 46: /cgm2_6/ptodata/1/paa/US602_COMB.pep:**
- 47: /cgm2_6/ptodata/1/paa/US603_COMB.pep:**
- 48: /cgm2_6/ptodata/1/paa/US604_COMB.pep:**
- 49: /cgm2_6/ptodata/1/paa/US605_COMB.pep:**
- 50: /cgm2_6/ptodata/1/paa/US606_COMB.pep:**
- 51: /cgm2_6/ptodata/1/paa/US607_COMB.pep:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	6909	100.0	1332	1	PCT-US03-22467-4	Sequence 4, Appli
2	6909	100.0	1332	1	PCT-US03-40113-4	Sequence 4, Appli
3	6909	100.0	1332	1	PCT-US05-01469-4	Sequence 4, Appli
4	6909	100.0	1332	36	US-10-621-758A-4	Sequence 4, Appli
5	6909	100.0	1332	36	US-10-646-301A-4	Sequence 4, Appli
6	6909	100.0	1332	36	US-10-663-208A-4	Sequence 4, Appli
7	6909	100.0	1332	37	US-10-736-769-4	Sequence 4, Appli
8	6909	100.0	1332	37	US-10-750-386-4	Sequence 4, Appli
9	6896	99.8	1332	1	PCT-US01-04098A-1831	Sequence 1831, Ap
10	6896	99.8	1332	32	US-10-239-316-9	Sequence 9, Appli
11	6896	99.8	1332	32	US-10-258-899A-1831	Sequence 1831, Ap
12	6896	99.8	1332	32	US-10-293-244-1831	Sequence 1831, Ap
13	6896	99.8	1332	42	US-11-218-141-1831	Sequence 1831, Ap
14	6896	99.8	1332	42	US-11-242-459-9	Sequence 9, Appli
15	6872.5	99.5	1359	1	PCT-US01-04098A-1830	Sequence 1830, Ap
16	6872.5	99.5	1359	1	PCT-US03-22467-44	Sequence 44, Appl
17	6872.5	99.5	1359	1	PCT-US03-40113-44	Sequence 44, Appl
18	6872.5	99.5	1359	1	PCT-US05-01469-44	Sequence 44, Appl
19	6872.5	99.5	1359	1	PCT-US05-27579-22	Sequence 22, Appl
20	6872.5	99.5	1359	32	US-10-258-899A-1830	Sequence 1830, Ap
21	6872.5	99.5	1359	32	US-10-293-244-1830	Sequence 1830, Ap
22	6872.5	99.5	1359	36	US-10-621-758A-44	Sequence 44, Appl
23	6872.5	99.5	1359	36	US-10-646-301A-44	Sequence 44, Appl
24	6872.5	99.5	1359	36	US-10-663-208A-44	Sequence 44, Appl
25	6872.5	99.5	1359	37	US-10-736-769-44	Sequence 44, Appl
26	6872.5	99.5	1359	37	US-10-750-386-44	Sequence 44, Appl
27	6872.5	99.5	1359	42	US-11-218-141-1830	Sequence 1830, Ap
28	6868.5	99.4	1359	31	US-10-170-205B-21686	Sequence 21686, A
29	6868.5	99.4	1359	48	US-60-453-050-8956	Sequence 8956, Ap
30	6868.5	99.4	1359	48	US-60-466-412-8956	Sequence 8956, Ap
31	6868.5	99.4	1359	48	US-60-466-412-8956	Sequence 8956, Ap
32	6536	94.6	1344	1	PCT-US01-08631-53052	Sequence 53052, A
33	6536	94.6	1344	34	US-10-450-763-53052	Sequence 53052, A
34	5421.5	78.5	1331	1	PCT-US03-22467-2	Sequence 2, Appli
35	5421.5	78.5	1331	1	PCT-US03-40113-2	Sequence 2, Appli
36	5421.5	78.5	1331	1	PCT-US05-01469-2	Sequence 2, Appli
37	5421.5	78.5	1331	36	US-10-621-758A-2	Sequence 2, Appli
38	5421.5	78.5	1331	36	US-10-646-301A-2	Sequence 2, Appli
39	5421.5	78.5	1331	36	US-10-663-208A-2	Sequence 2, Appli
40	5421.5	78.5	1331	37	US-10-736-769-2	Sequence 2, Appli
41	5421.5	78.5	1331	37	US-10-750-386-2	Sequence 2, Appli
42	5407	78.3	1333	1	PCT-US03-22467-12	Sequence 12, Appl
43	5407	78.3	1333	1	PCT-US03-40113-12	Sequence 12, Appl
44	5407	78.3	1333	1	PCT-US05-01469-12	Sequence 12, Appl
45	5407	78.3	1333	1	PCT-US05-27579-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
PCT-US03-22467-4
; Sequence 4, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation

; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JBO1603-K-WI			
; CURRENT APPLICATION NUMBER: PCT/US03/22467			
; CURRENT FILING DATE: 2003-07-17			
; NUMBER OF SEQ ID NOS: 50			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 4			
; LENGTH: 1332			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
; PCT-US03-40113-4			
Query Match	100.0%;	Score 6909;	DB 1; Length 1332;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 1332;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	MAEAGLRGWLWALLRLAQSPPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSPPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVSCLSN	60
QY	61	TPARKITGDHLLILOKICPRLYTGPNTOACCSAKQLVSLSEASLITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLLILOKICPRLYTGPNTOACCSAKQLVSLSEASLITKALLTRCPACSDNF	120
QY	121	VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRVPA	180
Db	121	VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRVPA	180
QY	181	ATLAVGTMCVGYGSSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240
Db	181	ATLAVGTMCVGYGSSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240
QY	241	ARCNEQGDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILCSFAVVTI	300
Db	241	ARCNEQGDVATCSQDCAACPAIARPOALDSTFYLGQMPGSLVLIILCSFAVVTI	300
QY	301	LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQMGWGWASWPLTILVLSV	360
Db	301	LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQMGWGWASWPLTILVLSV	360
QY	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHFQFFRTNQVILTPNRSSY	420
Db	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHFQFFRTNQVILTPNRSSY	420
QY	421	RYDSLILGPKNFSGIILDLILLLELLEQLERHLQWSPQAQRNLSLQICVAPLNPDT	480
Db	421	RYDSLILGPKNFSGIILDLILLLELLEQLERHLQWSPQAQRNLSLQICVAPLNPDT	480
QY	481	SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTKDGTALAL	540
Db	481	SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTKDGTALAL	540
QY	541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEEFLEEM	600
Db	541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEWEEFLEEM	600
QY	601	RAFQRRMAGMFQVTFPAAERSLDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV	660
Db	601	RAFQRRMAGMFQVTFPAAERSLDEINRTTAEDLPFATSYIVIFLYISLALGYSYSSWSRV	660
QY	661	MVDSKATLGLGGVAVLVGAVMAAGPFSYLGIRSSVLQWVPFLVLSVGADNIFIFVLE	720
Db	661	MVDSKATLGLGGVAVLVGAVMAAGPFSYLGIRSSVLQWVPFLVLSVGADNIFIFVLE	720
QY	721	YQRLPRPGPREVHIHGRALGRVAPSMMLCSLSEAIFFLGALTPMPAVRTFALTSGLA	780
Db	721	YQRLPRPGPREVHIHGRALGRVAPSMMLCSLSEAIFFLGALTPMPAVRTFALTSGLA	780
QY	781	ILDFLQWAFVALLSLDSKQEAERLDVCCVCKPQELPPQGGEGLLGLFFQKAYAPFL	840
Db	781	ILDFLQWAFVALLSLDSKQEAERLDVCCVCKPQELPPQGGEGLLGLFFQKAYAPFL	840
QY	841	LHWITRGVLLLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900

QY 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
QY 181 ATLAAGTMCVYGSALCNQRLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
DB 181 ATLAAGTMCVYGSALCNQRLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
QY 241 ARCNSQGDVATCSCQDCAASCPAIPALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNSQGDVATCSCQDCAASCPAIPALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKGTSLSDKLFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKGTSLSDKLFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFFRTNQTIVLTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFFRTNQTIVLTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSQDICYAPLNPNT 480
DB 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTKOGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTKOGTALAL 540
QY 541 SCHADYGAPVFPFLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCHADYGAPVFPFLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFORMAGMFWOTFAERSLEDEINRTTAEDLPFATSIVITFLYISALGYSWSRV 660
DB 601 RAFORMAGMFWOTFAERSLEDEINRTTAEDLPFATSIVITFLYISALGYSWSRV 660
QY 661 MYDSKATLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGDNIPIFVLE 720
DB 661 MYDSKATLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGDNIPIFVLE 720
QY 721 YQRLPRRPEPREVHIGRALGRVAPSMILCSISEAICFFLIGALTMPAVRTTALTSLGLAV 780
DB 721 YQRLPRRPEPREVHIGRALGRVAPSMILCSISEAICFFLIGALTMPAVRTTALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKREASRLDVCCVKPOELPPCGOEGLLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKREASRLDVCCVKPOELPPCGOEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSEAGNNAICSSAGCNGNFSTQKIQVATFPFGQSYLAIPASSWDDDFIDW 960
DB 901 YFVTTILGYNFSEAGNNAICSSAGCNGNFSTQKIQVATFPFGQSYLAIPASSWDDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKCPSVNSLNCIKCMSITMGSVRPSVEQFKHYPFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKCPSVNSLNCIKCMSITMGSVRPSVEQFKHYPFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMWYHKPLKNSQDVTYALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMWYHKPLKNSQDVTYALRAARELANITADL 1080
QY 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
DB 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGLL 1140
QY 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVAGVAMTNLPGILLVLGLAKAQLIQIPFFRLNLLITLGLLHGLVFLPV 1260

DB 1201 AKEATISMSGSAVAGVAMTNLPGILLVLGLAKAQLIQIPFFRLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDVNPALALQEKRAEBAVAVMVASCPNHPSTADNIYVNHSPFSGIKGAGA 1320
DB 1261 ILSYVGPDVNPALALQEKRAEBAVAVMVASCPNHPSTADNIYVNHSPFSGIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 3
PCT-US05-01469-4
; Sequence 4, Application PC/TUS0501469
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; TITLE OF INVENTION: NPC11 (NEC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF
; FILE REFERENCE: 36134-PCT 074669.0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US 60/537,341
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01469-4

Query Match 100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLSEASLITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLSEASLITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSNQSLEFINTRVAQLGAGQLPAVVAEAFYQHSFARQSDSCSRVRPAA 180
QY 181 ATLAAGTMCVYGSALCNQRLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
DB 181 ATLAAGTMCVYGSALCNQRLNFGDGTGNGLAPLDITFHLLEPCQAVSGIQLNEGV 240
QY 241 ARCNSQGDVATCSCQDCAASCPAIPALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNSQGDVATCSCQDCAASCPAIPALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKGTSLSDKLFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKGTSLSDKLFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFFRTNQTIVLTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPFFRTNQTIVLTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSQDICYAPLNPNT 480
DB 421 RYDSLILGPKNPSGILDLLELLELQERLHLQWSPEAQRNLSQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTKOGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDHFLYCANAPLTKOGTALAL 540
QY 541 SCHADYGAPVFPFLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600

```
Db 541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAQORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
Db 601 RAQORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGELLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGELLGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLEIPASSWVDDFIDW 960
Db 901 YFVTTILGNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLEIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
QY 1021 NTKCPKGGAAVSTSNLTSQVLAERMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NTKCPKGGAAVSTSNLTSQVLAERMAVHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEPFYTITNVFYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140
Db 1081 RKVPGTDPAFEPFYTITNVFYEQLTILPEGLFMLSCLVPTFAVSCILLGLDLSGILL 1140
QY 1141 NLLSIVMLIVDTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLER 1200
Db 1141 NLLSIVMLIVDTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLER 1200
QY 1201 AKEATISGMSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRNLNLTITLLGLHGLVFLPV 1260
Db 1201 AKEATISGMSAVPAGVAMTNLPGILVLGLAKAQLIQIPFRNLNLTITLLGLHGLVFLPV 1260
QY 1261 ILSYVGPDPNPALEQKRAEBAVAVMVASCENHPRSRVSTADNITYVNHSPSGSIKGAGA 1320
Db 1261 ILSYVGPDPNPALEQKRAEBAVAVMVASCENHPRSRVSTADNITYVNHSPSGSIKGAGA 1320
QY 1321 ISNFLPNNGROF 1332
Db 1321 ISNFLPNNGROF 1332
```

RESULT 4

```
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4
Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGSIAMTNSVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGSIAMTNSVCSLSN 60
QY 61 TPARKITGDHLILQKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILQKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAGVTGCGYGSALCNAQRLNPOGDTGNGLAPLDITFHLLERPGQAVGSGIQPLNEGV 240
Db 181 ATLAGVTGCGYGSALCNAQRLNPOGDTGNGLAPLDITFHLLERPGQAVGSGIQPLNEGV 240
QY 241 ARCHESQDDVATSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCHESQDDVATSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTPVELVSNAPNSQARSEKAFHQHGFPPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTPVELVSNAPNSQARSEKAFHQHGFPPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNPGSLDLDLLELEQLERLHLQVMSPEAQNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPGSLDLDLLELEQLERLHLQVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCINSLLYQFQNNRTLLLTANTQTMGOTSQVMDKHPLYCANAPLTFKDGDTALAL 540
Db 481 SLYDCINSLLYQFQNNRTLLLTANTQTMGOTSQVMDKHPLYCANAPLTFKDGDTALAL 540
QY 541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCWADYGAVPFPFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAQORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
Db 601 RAQORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGADNIFIVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGELLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGELLGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLEIPASSWVDDFIDW 960
Db 901 YFVTTILGNFSSEAGNNAICSSAGCNFSFTQKIYATEFFPQSYLEIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQFHXYLPWFINDRP 1020
```

```
Db      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGSVRPSVQFHKYLPWFNDRP 1020
Qy      1021  NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANTADL 1080
Db      1021  NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANTADL 1080
Qy      1081  RKVPGTDPAFEVPPYITINVFYBQYLTIPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db      1081  RKVPGTDPAFEVPPYITINVFYBQYLTIPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy      1141  NLLSIVMLVDVTGFMALWDISYNAVSLNLNSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db      1141  NLLSIVMLVDVTGFMALWDISYNAVSLNLNSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Qy      1201  AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Db      1201  AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Qy      1261  ILSYVGPVDVNPALALQKRAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db      1261  ILSYVGPVDVNPALALQKRAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Qy      1321  ISNFLPNNGRQF 1332
Db      1321  ISNFLPNNGRQF 1332

RESULT 5
US-10-646-301A-4
; Sequence 4, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4

Query Match      100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MAEAGLRGWLWALLRLAQSEPYTTHOPGYCAFYDECKNPELSGSLMTLSNVCSLSN 60
Db      1  MAEAGLRGWLWALLRLAQSEPYTTHOPGYCAFYDECKNPELSGSLMTLSNVCSLSN 60
Qy      61  TPARKITGDHLLILLOKICPRLVTGPNTOACCSAKQLVSLBASLITKALLTRCPACSDNF 120
Db      61  TPARKITGDHLLILLOKICPRLVTGPNTOACCSAKQLVSLBASLITKALLTRCPACSDNF 120
Qy      121  VNLHCHNTCSPNOSLFINVTRVAQIGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
Db      121  VNLHCHNTCSPNOSLFINVTRVAQIGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
Qy      181  ATLVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db      181  ATLVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy      241  ARCNSQGDVDVATSCQDCAASCPAIPALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
```

```
Db      241  ARCNSQGDVDVATSCQDCAASCPAIPALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy      301  LLVGFVAPARDKSMVDKPKGTSLSKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db      301  LLVGFVAPARDKSMVDKPKGTSLSKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Qy      361  IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTPNRSY 420
Db      361  IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTPNRSY 420
Qy      421  RYDLSLLGPKNSGILDLDDLLELLELQRLHRLQWSPEAQRNISLQDICYAPLNPNDT 480
Db      421  RYDLSLLGPKNSGILDLDDLLELLELQRLHRLQWSPEAQRNISLQDICYAPLNPNDT 480
Qy      481  SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDMKDHFLYCANAPLTKDGTALAL 540
Db      481  SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDMKDHFLYCANAPLTKDGTALAL 540
Qy      541  SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLMEAEFLEEM 600
Db      541  SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLMEAEFLEEM 600
Qy      601  RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIIVIFLYISLALSGYSWSRV 660
Db      601  RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIIVIFLYISLALSGYSWSRV 660
Qy      661  MYDSKATIGLGGVAVVLGAVMAAMGPFSSYLGRSSIVILQVVPFLVLSVGAADNIFIVLE 720
Db      661  MYDSKATIGLGGVAVVLGAVMAAMGPFSSYLGRSSIVILQVVPFLVLSVGAADNIFIVLE 720
Qy      721  YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEAI CFFELGALTMPAVRTALTSLGLAV 780
Db      721  YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEAI CFFELGALTMPAVRTALTSLGLAV 780
Qy      781  ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGELLGFFQKAYAPFL 840
Db      781  ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQEGELLGFFQKAYAPFL 840
Qy      841  LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
Db      841  LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYLFLNRYFEVGAPV 900
Qy      901  YFVTLTYNFSSEAGMNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDIDW 960
Db      901  YFVTLTYNFSSEAGMNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDIDW 960
Qy      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGSVRPSVQFHKYLPWFNDRP 1020
Db      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGSVRPSVQFHKYLPWFNDRP 1020
Qy      1021  NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANTADL 1080
Db      1021  NIKCPKGLAAYSTSVNLTSQGVLASRFMAHYHKPLKNSQDYTEALRAARELAANTADL 1080
Qy      1081  RKVPGTDPAFEVPPYITINVFYBQYLTIPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db      1081  RKVPGTDPAFEVPPYITINVFYBQYLTIPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy      1141  NLLSIVMLVDVTGFMALWDISYNAVSLNLNSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db      1141  NLLSIVMLVDVTGFMALWDISYNAVSLNLNSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Qy      1201  AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Db      1201  AKEATISGMSAVPAGVAMTNLPGLVGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Qy      1261  ILSYVGPVDVNPALALQKRAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db      1261  ILSYVGPVDVNPALALQKRAEAAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Qy      1321  ISNFLPNNGRQF 1332
Db      1321  ISNFLPNNGRQF 1332
```



```
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-4

Query Match      100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGHWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGHWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSFLINVTVAQAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
DB 121 VNLHCHNTCSPNQSFLINVTVAQAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
QY 181 ATAVGTGTCVGVYSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATAVGTGTCVGVYSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQGDVATCSCDCAACPAIARPOALDSTFYLGOMPGSLVLIILCSVAVVTI 300
DB 241 ARCNESQGDVATCSCDCAACPAIARPOALDSTFYLGOMPGSLVLIILCSVAVVTI 300
QY 301 LLVGRFVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWVASWPLTLVLVS 360
DB 301 LLVGRFVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWVASWPLTLVLVS 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICVAPLNPDNT 480
DB 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
QY 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAPLEEM 600
DB 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAPLEEM 600
QY 601 RAPORMAGMFOVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV 660
DB 601 RAPORMAGMFOVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV 660
QY 661 WYDSKATILGCGVAVVLGAVMAAGMFPFVSLGRSSILVILQVPELVLSGADNIFVLE 720
DB 661 WYDSKATILGCGVAVVLGAVMAAGMFPFVSLGRSSILVILQVPELVLSGADNIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEACFFLGLATPMPAVRTPALTSLGLAV 780
DB 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEACFFLGLATPMPAVRTPALTSLGLAV 780
QY 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQSGELLGFPQKAYAPPL 840
DB 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQSGELLGFPQKAYAPPL 840
QY 841 LHWITRGVVLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYLFLNRYFEVGAPV 900
QY 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDVDFIDW 960
DB 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDVDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKPKGGLAAYSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKPKGGLAAYSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYTIITNVFYEQYLTILPEGLFMLSICLVPTFAVSCLLGLDLSGLL 1140
DB 1081 RKVPGTDPAPFVPPYTIITNVFYEQYLTILPEGLFMLSICLVPTFAVSCLLGLDLSGLL 1140
QY 1141 NLLSVIMLIVDVTGFMALWDISYNAVSLINLSVAGMSVEFYSHITRSPAISTKPTWLER 1200
DB 1141 NLLSVIMLIVDVTGFMALWDISYNAVSLINLSVAGMSVEFYSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGIILVLGLAKAQLIQIFFRLNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVFAGVAMTNLPGIILVLGLAKAQLIQIFFRLNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRABEAVAAVMVASCPNHPSPRSTADNIYVNHSPGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRABEAVAAVMVASCPNHPSPRSTADNIYVNHSPGSIKGAGA 1320
QY 1321 ISNPLPNNGRGP 1332
DB 1321 ISNPLPNNGRGP 1332

RESULT 8
US-10-750-386-4
; Sequence 4, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujjainwalla, Feroze
; APPLICANT: Altman, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONIST
; FILE REFERENCE: A36104 074669 0103
; CURRENT APPLICATION NUMBER: US/10750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-750-386-4

Query Match      100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLRGHWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGHWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSFLINVTVAQAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
DB 121 VNLHCHNTCSPNQSFLINVTVAQAGOLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
QY 181 ATAVGTGTCVGVYSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATAVGTGTCVGVYSALCNAQRWLNFGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQGDVATCSCDCAACPAIARPOALDSTFYLGOMPGSLVLIILCSVAVVTI 300
DB 241 ARCNESQGDVATCSCDCAACPAIARPOALDSTFYLGOMPGSLVLIILCSVAVVTI 300
QY 301 LLVGRFVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWVASWPLTLVLVS 360
DB 301 LLVGRFVAPARDKRWDPKGTSTLSKLSFSTHTLLGQFQCGWTWVASWPLTLVLVS 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFPHDQHPFPRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICVAPLNPDNT 480
DB 421 RYDSLILGPKNPSGILDLLELELQERLHLQVNSPEAQRNLSIODICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLPCANAPLTKDGTAL 540
QY 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAPLEEM 600
DB 541 SCWADYGAPVFPFLAIGGKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAPLEEM 600
QY 601 RAPORMAGMFOVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV 660
DB 601 RAPORMAGMFOVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV 660
QY 661 WYDSKATILGCGVAVVLGAVMAAGMFPFVSLGRSSILVILQVPELVLSGADNIFVLE 720
DB 661 WYDSKATILGCGVAVVLGAVMAAGMFPFVSLGRSSILVILQVPELVLSGADNIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEACFFLGLATPMPAVRTPALTSLGLAV 780
DB 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEACFFLGLATPMPAVRTPALTSLGLAV 780
QY 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQSGELLGFPQKAYAPPL 840
DB 781 ILDFLLQWSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQSGELLGFPQKAYAPPL 840
QY 841 LHWITRGVVLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYLFLNRYFEVGAPV 900
QY 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDVDFIDW 960
DB 901 YFVTTLTGYNFSSEAGNNAICSSAGCWNFSFTQKIQVATEFPQSYLAIPASSWVDVDFIDW 960
```

Db	181	ATLAVGTGCGVGSALCNAQRWLNFGDITGNGLAPLDITFHLLPEQAVGSGIQLPNEG	240
Qy	241	ARCNSQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300
Db	241	ARCNSQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300
Qy	301	LLVGFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY	420
Db	361	IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY	420
Qy	421	RYDSLILGPNKFSIGILDLLLELLELQBLRLHLQWSPEAQRNLSQDICYAPLNPDNT	480
Db	421	RYDSLILGPNKFSIGILDLLLELLELQBLRLHLQWSPEAQRNLSQDICYAPLNPDNT	480
Qy	481	SLYDCCINSILQYFQNNRTLLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTIKDGTAL	540
Db	481	SLYDCCINSILQYFQNNRTLLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTIKDGTAL	540
Qy	541	SCWADYGAPVFPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEFALEEM	600
Db	541	SCWADYGAPVFPLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEFALEEM	600
Qy	601	RAFQRMAGMFQVTFPQRSLEDEINRTTAEDLPATPSYIVIFLYISALGSYSWSRV	660
Db	601	RAFQRMAGMFQVTFPQRSLEDEINRTTAEDLPATPSYIVIFLYISALGSYSWSRV	660
Qy	661	MVDSKATLGGVAVVLGAVMAAMGPFSLGIRSSILVILQVVPFLVLSVGADNIFVLE	720
Db	661	MVDSKATLGGVAVVLGAVMAAMGPFSLGIRSSILVILQVVPFLVLSVGADNIFVLE	720
Qy	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFELGALTMPAPVTRTALTSGLAV	780
Db	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEALICFFELGALTMPAPVTRTALTSGLAV	780
Qy	781	ILDFLQMSAFVALLSLDSKROBASRLDVCCCKVQOELPPPGQEGELLIGFPQKAYAPPL	840
Db	781	ILDFLQMSAFVALLSLDSKROBASRLDVCCCKVQOELPPPGQEGELLIGFPQKAYAPPL	840
Qy	841	LHWITRGVLLFLALFGVLSYMSCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVLSYMSCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAPV	900
Qy	901	YFVTTILGYNFSSEAGMNAICSSAGCNFSTQKI QYATEFPQSYLAI PASSWVDDFIDW	960
Db	901	YFVTTILGYNFSSEAGMNAICSSAGCNFSTQKI QYATEFPQSYLAI PASSWVDDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNMSITMGSVRPSVEQHKYLPWFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNMSITMGSVRPSVEQHKYLPWFLNDRP	1020
Qy	1021	NIKCPKGLAAYSTSVNLSDGOVLASRFMAHYKPLKNSODYTEALRAARELAANTADL	1080
Db	1021	NIKCPKGLAAYSTSVNLSDGOVLASRFMAHYKPLKNSODYTEALRAARELAANTADL	1080
Qy	1081	RKVPGTDPAPFVPPYITINNVFEQYLITLPEGLFMSLCLVPTFAVSCILLGLDLSRGLL	1140
Db	1081	RKVPGTDPAPFVPPYITINNVFEQYLITLPEGLFMSLCLVPTFAVSCILLGLDLSRGLL	1140
Qy	1141	NLLSIYMLIVDTVGFMAWHDISYNAVSLNLINLSAVGMSVEFVSHITRSPAISTKTWLER	1200
Db	1141	NLLSIYMLIVDTVGFMAWHDISYNAVSLNLINLSAVGMSVEFVSHITRSPAISTKTWLER	1200
Qy	1201	AKEATISMGSAVAGVAMTNLPGLVGLAKAQLIQIFFFRNLNLTLLGLLHGLVFLPV	1260
Db	1201	AKEATISMGSAVAGVAMTNLPGLVGLAKAQLIQIFFFRNLNLTLLGLLHGLVFLPV	1260
Qy	1261	ILSYGPDVNPALALEQKRAEBAVAAMVASCNPHPRSRVSTADNIYVNHSPFGSIKGAGA	1320

Db	1261	ILSYGPDVNPALALEQKRAEBAVAAMVASCNPHPRSRVSTADNIYVNHSPFGSIKGAGA	1320
Qy	1321	ISNFLPNNGRQF 1332	
Db	1321	ISNFLPNNGRQF 1332	
RESULT 9			
PCT-US01-04098A-1831			
; Sequence 1831, Application PC/TUS0104098A			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc.			
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides			
; FILE REFERENCE: 21272-029			
; CURRENT APPLICATION NUMBER: PCT/US01/04098A			
; CURRENT FILING DATE: 2001-02-05			
; PRIOR APPLICATION NUMBER: Not Yet Assigned			
; PRIOR FILING DATE: 2001-01-30			
; PRIOR APPLICATION NUMBER: 09/728,422			
; PRIOR FILING DATE: 2000-11-30			
; PRIOR APPLICATION NUMBER: 09/693,325			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR APPLICATION NUMBER: 09/663,561			
; PRIOR FILING DATE: 2000-09-15			
; PRIOR APPLICATION NUMBER: 09/654,936			
; PRIOR FILING DATE: 2000-09-01			
; PRIOR APPLICATION NUMBER: 09/620,325			
; PRIOR FILING DATE: 2000-07-19			
; PRIOR APPLICATION NUMBER: 09/598,075			
; PRIOR FILING DATE: 2000-06-20			
; PRIOR APPLICATION NUMBER: 09/560,875			
; PRIOR FILING DATE: 2000-04-27			
; PRIOR APPLICATION NUMBER: 09/496,914			
; PRIOR FILING DATE: 2000-02-03			
; NUMBER OF SEQ ID NOS: 3960			
; SOFTWARE: Custom			
; SEQ ID NO 1831			
; LENGTH: 1332			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
PCT-US01-04098A-1831			
Query Match 99.8%; Score 6896; DB 1; Length 1332;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPELSGLMTLSNVSCLSN	60
Qy	61	TPARKITGDHLLILQKICPRLYTGTNTQACCSAKQLVSLASLSITKALLTRCPACSNF	120
Db	61	TPARKITGDHLLILQKICPRLYTGTNTQACCSAKQLVSLASLSITKALLTRCPACSNF	120
Qy	121	VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAEYFQHSFAEQSYDSCSRVRVPA	180
Db	121	VNLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAEYFQHSFAEQSYDSCSRVRVPA	180
Qy	181	ATLAVGTWCVYGSALCNAQRWLNFGDITGNGLAPLDITFHLLPEQAVGSGIQLPNEG	240
Db	181	ATLAVGTWCVYGSALCNAQRWLNFGDITGNGLAPLDITFHLLPEQAVGSGIQLPNEG	240
Qy	241	ARCNSQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300
Db	241	ARCNSQGDVATCSQDCODCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI	300
Qy	301	LLVGFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Db	301	LLVGFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV	360
Qy	361	IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY	420
Db	361	IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY	420

Qy	421	RYDSLLGPKNFSGILDLIDLLLELELELELOERLRLHLOWSPQAQRNISILODICYAPLNDPNT	480
Db	421	RYDSLLGPKNFSGILDLLELELELELOERLRLHLOWSPQAQRNISILODICYAPLNDPNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKOHFLYCANAPITFKDGTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKOHFLYCANAPITFKDGTALAL	540
Qy	541	SCMADYGAPVPFLAIAGYKGKQYSEAEALIMTFSLNYPAGDPRLAQAKLWBEAFLEEM	600
Db	541	SCMADYGAPVPFLAIAGYKGKQYSEAEALIMTFSLNYPAGDPRLAQAKLWBEAFLEEM	600
Qy	601	RAFORRMAGMPQVTTAERSLEDBINRTTAEDLPFATSYIVIFLYTSLALGSYSSWSRV	660
Db	601	RAFORRMAGMPQVTTAERSLEDBINRTTAEDLPFATSYIVIFLYTSLALGSYSSWSRV	660
Qy	661	MVDSKATLIGGVAVVLGAVMAAGPFSYLGIRSSVLIVQVVPFLVSVGADNITFIFVLE	720
Db	661	MVDSKATLIGGVAVVLGAVMAAGPFSYLGIRSSVLIVQVVPFLVSVGADNITFIFVLE	720
Qy	721	YQRLPRPGPREVHI GRALGRVAPSMLLCSLSEAI CFFLCALTPMPAVRTFALTSGLAV	780
Db	721	YQRLPRPGPREVHI GRALGRVAPSMLLCSLSEAI CFFLCALTPMPAVRTFALTSGLAV	780
Qy	781	ILDFLQWSAFVALLSDSKRQEASRLDVCCKVPQELPPPGQEGEGLLGFQKAYAPFL	840
Db	781	ILDFLQWSAFVALLSDSKRQEASRLDVCCKVPQELPPPGQEGEGLLGFQKAYAPFL	840
Qy	841	LHWITRGVVLVLLFIALFGVSLYSWCHISVGLDQBELAPKDSYLLDYFLFNRYPEVGAPV	900
Db	841	LHWITRGVVLVLLFIALFGVSLYSWCHISVGLDQBELAPKDSYLLDYFLFNRYPEVGAPV	900
Qy	901	YFVTTLCYNFSSEAGMNAICSSAGCINNFSFTOKI QYATEPQOSYLAI PASSWVDDFDW	960
Db	901	YFVTTLCYNFSSEAGMNAICSSAGCINNFSFTOKI QYATEPQOSYLAI PASSWVDDFDW	960
Qy	961	LTSSCCRLYISGNPKDPCPSTVNSLNCNLKNCMSITWGSVRPSPVEQFHYLPWFLNDRP	1020
Db	961	LTSSCCRLYISGNPKDPCPSTVNSLNCNLKNCMSITWGSVRPSPVEQFHYLPWFLNDRP	1020
Qy	1021	NIKCPKGLAAYSTSVNLTSDQVLA SRFMA YHKPLKNSQDYTEALRAARELANITADL	1080
Db	1021	NIKCPKGLAAYSTSVNLTSDQVLA SRFMA YHKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAPFVFPYTTITNPFYEOYLITLPEGLFMLSICLVPTPAVSCLLGLDLRSGLL	1140
Db	1081	RKVPGTDPAPFVFPYTTITNPFYEOYLITLPEGLFMLSICLVPTPAVSCLLGLDLRSGLL	1140
Qy	1141	NLLSIVMLVDVTGFMALWDISYNAVSLINLVSAVGMSVFVSHITSPAI STKPTWLER	1200
Db	1141	NLLSIVMLVDVTGFMALWDISYNAVSLINLVSAVGMSVFVSHITSPAI STKPTWLER	1200
Qy	1201	AKRATISWGSVAVFAGVANTMLPGILVLGLAKAQI QIFFFLRNLIIITLLGLLHGLVFLPV	1260
Db	1201	AKRATISWGSVAVFAGVANTMLPGILVLGLAKAQI QIFFFLRNLIIITLLGLLHGLVFLPV	1260
Qy	1261	ILSYVGPDPVNPALAEQKRAEEAVAAVMVASCPNHPSRVSTADNIYVNHSPFEGSIKGAGA	1320
Db	1261	ILSYVGPDPVNPALAEQKRAEEAVAAVMVASCPNHPSRVSTADNIYVNHSPFEGSIKGAGA	1320
Qy	1321	ISNPLPNNGRQF 1332	
Db	1321	ISNPLPNNGRQF 1332	

RESULT 10

RESULT 10
US-10-239-316-9

US-10-239-316-9
; Sequence 9, Application US/10239316

GENERAL INFORMATION:

APPLICANT: TANIYAMA, Yoshio

APPLICANT: KITA, Shunbun

APPLICANT: SATOMI, Tomoko Komiyama

```

, TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
, FILE REFERENCE: 2703US0P
, CURRENT APPLICATION NUMBER: US/10/239,316
, PRIOR FILING DATE: 2002-09-19
, PRIOR APPLICATION NUMBER: PCT/JP01/02279
, PRIOR FILING DATE: 2001-03-22
, PRIOR APPLICATION NUMBER: JP2000-088595
, PRIOR FILING DATE: 2000-03-24
, NUMBER OF SEQ ID NOS: 59
, SEQ ID NO 9
, LENGTH: 1332
, TYPE: PRT
, ORGANISM: Human
US-10-239-316-9

```

```
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKPQELPPGQCBGLLLGFFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPV 900
QY 901 YFVTTLGNPSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSVLAIPASSWVDDFDW 960
Db 901 YFVTTLGNPSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSVLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNLCNMSITMGSVRPSVEQFHKYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCSTVNSLNCNLCNMSITMGSVRPSVEQFHKYLPWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVGTDPAFEPFPPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVGTDPAFEPFPPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSVMLILVDVGVFMALWDISYNAVSLINLYSAGMSVEFVSHITRSPFAISTKPTWLER 1200
Db 1141 NLLSVMLILVDVGVFMALWDISYNAVSLINLYSAGMSVEFVSHITRSPFAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGLILVLGLAKAQLTOIPFFRLNLITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGLILVLGLAKAQLTOIPFFRLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVNPALALEOKBAEAAVAVMVASCPNHPSPRSTADNLYVNHSPSGSIKGAGA 1320
Db 1261 ILSYVGPVNPALALEOKBAEAAVAVMVASCPNHPSPRSTADNLYVNHSPSGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 11
US-10-258-899A-1831
; Sequence 1831. Application US/10258899A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunqueing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258, 899A
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
```

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-258-899A-1831

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHILLQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILLQKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSPNQSIFINVTVAQGLGQLPVAVAYEAFYOHSPAEOQSDYSCSRVRVPA 180
Db 121 VNLCHNTCSPNQSIFINVTVAQGLGQLPVAVAYEAFYOHSPAEOQSDYSCSRVRVPA 180
QY 181 ATLAGVTGCVGVGSALCNAQRLNFGQDGTGNGLAPLDITTFHLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAGVTGCVGVGSALCNAQRLNFGQDGTGNGLAPLDITTFHLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQGDVATCSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSCQDCAASCPAIAAPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVPKGTSLSDKLSFSTHTLLGQFCQGWCTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVPKGTSLSDKLSFSTHTLLGQFCQGWCTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTAPNRSY 420
QY 421 RYDSLGLGPKNPSGILDLDLLELLELRLHQLVMSPEAQRNLSQDICYAPLNPDT 480
Db 421 RYDSLGLGPKNPSGILDLDLLELLELRLHQLVMSPEAQRNLSQDICYAPLNPDT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDFLYCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVMDKDFLYCANAPLTKDGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFOVTTFAERSLEDEINRTAEDLPITPASYIIVIFLYISALGYSYWSRV 660
Db 601 RAFQRMAGMFOVTTFAERSLEDEINRTAEDLPITPASYIIVIFLYISALGYSYWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSVLGRSSILVILQVVPFLVLSGADNIFIFVLE 720
```

```
QY 721 YORLPBPGEPREVHIGRALGRVAPSMILCSLSEAI CFFLGALTMPAVRTALTSGLAV 780
Db 721 YORLPBPGEPREVHIGRALGRVAPSMILCSLSEAI CFFLGALTMPAVRTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPPGQEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKI QYATEFPQSYLAIPASSWVDDPIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKI QYATEFPQSYLAIPASSWVDDPIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCCKMCSITMGSVPSVEQFHKYLIPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCCKMCSITMGSVPSVEQFHKYLIPWFLNDRP 1020
QY 1021 NIKCPKGLAAVSTSVNLTSQVTLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAVSTSVNLTSQVTLASRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFVFPPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFVFPPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSTVMTILVDVTGFMALWDISYNAVSLNLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSTVMTILVDVTGFMALWDISYNAVSLNLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFFRNLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332
```

RESULT 12

```
US-10-293-244-1831
; Sequence 1831, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
```

```
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1831
```

```
Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLRLAQSPYTTIHOPGYCAFYDECGKNPELGGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLRLAQSPYTTIHOPGYCAFYDECGKNPELGGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICPRLVTGNTQACCSAKOLVLEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLVTGNTQACCSAKOLVLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSFLINVRVAQLGAGOLPAVVAEYAFYQHSFARQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSPNQSFLINVRVAQLGAGOLPAVVAEYAFYQHSFARQSYDSCSRVRVPA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLPEQQA VSGIQLNEG 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGQDGTNGGLAPLDITFHLLPEQQA VSGIQLNEG 240
QY 241 ARCNSQGGDVATCSCQDCNAASCPAIPARQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDVATCSCQDCNAASCPAIPARQALDSTFYLGMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNVDPKKTSLSDKLSFSTHTLLGQFFQCGWTWASWPLTILVLSV 360
Db 301 LLVGFRVAPARDKSNVDPKKTSLSDKLSFSTHTLLGQFFQCGWTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFETLTPDVELWSPNSQARSEKAFHDQHPFPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFETLTPDVELWSPNSQARSEKAFHDQHPFPFRTNQVILTAPNRSY 420
QY 421 RYDSLILGPKNPSGILDLDDLELELELELELELELELELELELELELELELELELE 480
Db 421 RYDSLILGPKNPSGILDLDDLELELELELELELELELELELELELELELELELELE 480
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDWKDHFYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOVDWKDHFYCANAPLTFKDGFTALAL 540
QY 541 SCMDYGA VFPFLAIGGYKGDYSEAEALINTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCMDYGA VFPFLAIGGYKGDYSEAEALINTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAPQRMAGMFQVTPAERSLEDEINRTTAEADLPFATSYIVIFLYISALSGYSWSRV 660
Db 601 RAPQRMAGMFQVTPAERSLEDEINRTTAEADLPFATSYIVIFLYISALSGYSWSRV 660
QY 661 MYDSKATILGCGVAVVLGAVMAAMGFPSVLGTRSSLVILQVVPVLVSVGAONIIFVLE 720
Db 661 MYDSKATILGCGVAVVLGAVMAAMGFPSVLGTRSSLVILQVVPVLVSVGAONIIFVLE 720
QY 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEAI CFFLGALTMPAVRTALTSGLAV 780
Db 721 YQRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEAI CFFLGALTMPAVRTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPPGQEGLLGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCCKVQDELPPPGQEGLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNMFSTQKI QYATEFPQSYLAIPASSWVDDPIDW 960
```

Db 901 YFVTTGLYNFSEAGNNAICSSAGCNFSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCSTVNSLNCCLKNCMSITMGSVRPSVEQPHKYLWPLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCSTVNSLNCCLKNCMSITMGSVRPSVEQPHKYLWPLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSVIMILVDVTFGMALWDISYNAVSLINLVSAGMSVEFVSHITRFAISTKPTWLER 1200
Db 1141 NLLSVIMILVDVTFGMALWGISYNAVSLINLVSAGMSVEFVSHITRFAISTKPTWLER 1200
Qy 1201 AKEATISGMSAVFAGVAMTNLPGILVLGLAKAQLIOIPFRLNMLITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGILVLGLAKAQLIOIPFRLNMLITLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPDPNPALALEOKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPNPALALEOKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 13
US-11-218-141-1831
; Sequence 1831, Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Yang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1831

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLAQSSEPYTTIHQPGYCAFYDECCKNPELSGSLMTLSNVCLSN 60

Db 1 MAEAGLRGWLWALLRLAQSSEPYTTIHQPGYCAFYDECCKNPELSGSLMTLSNVCLSN 60
Qy 61 TPARKITGDHILILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSNF 120
Db 61 TPARKITGDHILILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSNF 120
Qy 121 VNLCHNTCSPNQSLFNTVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLCHNTCSPNQSLFNTVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAVGTCMGVYGALCNQARWLNFQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTCMGVYGALCNQARWLNFQDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCHESQDDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300
Db 241 ARCHESQDDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIIILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSPSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSMVDPKKGTSLSDKLSPSTHTLLGQFFQGMGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQDHQFPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQDHQFPFRTNQVILTAPNRSY 420
Qy 421 RYDSLILGPKPFGSLDLDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
Db 421 RYDSLILGPKPFGSLDLDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 480
Qy 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGTTALAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGTTALAL 540
Qy 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLSEM 600
Db 541 SCMADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKWEAEFLSEM 600
Qy 601 RAFQRMAGMQVTFTAERSLEDEINRTAEDLPFATSYIVIFLYISIALGYSWSVRV 660
Db 601 RAFQRMAGMQVTFTAERSLEDEINRTAEDLPFATSYIVIFLYISIALGYSWSVRV 660
Qy 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVLQVVPFLVLSVADNIFIFVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAMGFFSYLGRSSILVLQVVPFLVLSVADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTTPMPAVRTALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFPLGALTTPMPAVRTALTSLGLAV 780
Qy 781 ILDFELLOMSAFVALLSLDSKROEASRLDVCCCKPQELPPQGGEGLLGFPQKAYAPFL 840
Db 781 ILDFELLOMSAFVALLSLDSKROEASRLDVCCCKPQELPPQGGEGLLGFPQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTGLYNFSEAGNNAICSSAGCNNFSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Db 901 YFVTTGLYNFSEAGNNAICSSAGCNNFSFTQKIQYATEFFEQSVLAIPASSWDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDFCSTVNSLNCCLKNCMSITMGSVRPSVEQPHKYLWPLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCSTVNSLNCCLKNCMSITMGSVRPSVEQPHKYLWPLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITNVFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140

Db 1081 RKVPGTDPAPFVPPYTTITNVFYEQYLTILPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGFMAWMDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWMDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAVPAVAVMTNLPGLVGLAKAQLIQIIPFRMLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVPAVAVMTNLPGLVGLAKAQLIQIIPFRMLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALQKRAEBAVAAVMVASCNPHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVNPALALQKRAEBAVAAVMVASCNPHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
RESULT 14
US-11-242-459-9
; Sequence 9, Application US/11242459
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9
Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MAEAGLRGMLLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGMLLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLILLOKICRPLYTGPNTOACCSAKQLVSLSEASITKALLTRCPACSDNP 120
Db 61 TPARKITGDHLLILLOKICRPLYTGPNTOACCSAKQLVSLSEASITKALLTRCPACSDNP 120
QY 121 VNLHCHTCSNOSLEFINTVRAQLGAGQLPAVVAEAFQHSFABQSDYSCSRVRVPA 180
Db 121 VNLHCHTCSNOSLEFINTVRAQLGAGQLPAVVAEAFQHSFABQSDYSCSRVRVPA 180
QY 181 ATLAVGTMCVGYGSALCNARWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQIPLNEGV 240
Db 181 ATLAVGTMCVGYGSALCNARWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQIPLNEGV 240
QY 241 ARCNEQGDVATCSQDCQCAAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCQCAAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSFTHLLGQFPQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKGTSLSDKLSFSFTHLLGQFPQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFPFRTNQVILTAPNRSY 420

Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFPGFPFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLLELLELQERLRLHQLQWSPQARNISLQIDICYAPLNPNT 480
Db 421 RYDLSLLGPKNPSGILDLLELLELQERLRLHQLQWSPQARNISLQIDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGKTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTPKDGKTALAL 540
QY 541 SCHADYGAVPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAALWEAELEEM 600
Db 541 SCHADYGAVPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAALWEAELEEM 600
QY 601 RAFORMMAGMPOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISLALGSYSWSRV 660
Db 601 RAFORMMAGMPOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEATCPFLGALTTPMPAVRTFALTSGLA 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEATCPFLGALTTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPGQEGLLGFFOKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCVKPQELPPGQEGLLGFFOKAYAPFL 840
QY 841 LEWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVEGAPV 900
Db 841 LEWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVEGAPV 900
QY 901 YFVTTILGYNPSSBAGNAICSSAGCNPFTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNPSSBAGNAICSSAGCNPFTQKIQYATEFPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQGLASRFMAYHKLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTVNLTSQGLASRFMAYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVPPYTTITNVFYEQYLTILPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAPFVPPYTTITNVFYEQYLTILPEGLFMLSLCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGFMAWMDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWMDISYNAVSLINLVSAVGMSEVFSVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAVPAVAVMTNLPGLVGLAKAQLIQIIPFRMLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVPAVAVMTNLPGLVGLAKAQLIQIIPFRMLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALQKRAEBAVAAVMVASCNPHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVNPALALQKRAEBAVAAVMVASCNPHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 15
PCT-US01-04098A-1830
; Sequence 1830, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

```
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; TYPE: PRT
; LENGTH: 1359
; ORGANISM: Homo sapiens
PCT-US01-04098A-1830

Query Match          99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy      Db      1  MAEAGLGRWLLALLRLAQSEPYTHIQPGCAFYDECGKPELSGSLMTLSNVCSLSN 60
Db      1  MAEAGLGRWLLALLRLAQSEPYTHIQPGCAFYDECGKPELSGSLMTLSNVCSLSN 60
Qy      61  TPARKITGDHLLILLQKICRPLYTGPNTQACCSAKQLVLSLEASISITKALLITRCPACSNF 120
Db      61  TPARKITGDHLLILLQKICRPLYTGPNTQACCSAKQLVLSLEASISITKALLITRCPACSNF 120
Qy      121  VNLHCHNTCSNQLFINTRVAQLGAGOLPAVVAEAFVQHSFABQSDSCSRVRVPA 180
Db      121  VNLHCHNTCSNQLFINTRVAQLGAGOLPAVVAEAFVQHSFABQSDSCSRVRVPA 180
Qy      181  ATLAVGTMCVGSALCNQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db      181  ATLAVGTMCVGSALCNQRWLNFGQDTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Qy      241  ARCNESQDDVATCSCQDCAACSPAIAARPOALDSTFYLGMPGSLVLIILCSVFAVVTI 300
Db      241  ARCNESQDDVATCSCQDCAACSPAIAARPOALDSTFYLGMPGSLVLIILCSVFAVVTI 300
Qy      301  LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db      301  LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy      361  IPVVALAAGLVFTLETTDPPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
Db      361  IPVVALAAGLVFTLETTDPPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
Qy      421  RYDSSLILGPKVFGSILDLDLLELLELOERLRLHQLVNSPEAQRNISLQDICYAPLNPNDT 480
Db      421  RYDSSLILGPKVFGSILDLDLLELLELOERLRLHQLVNSPEAQRNISLQDICYAPLNPNDT 480
Qy      481  SLVDCCINSLLQVFNQNRITLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKQGTALAL 540
Db      481  SLVDCCINSLLQVFNQNRITLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKQGTALAL 540
Qy      541  SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLLNNYPAGDPRLAQAKLWBEAFLEEM 600
Db      541  SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLLNNYPAGDPRLAQAKLWBEAFLEEM 600
```

```
Qy      601  RAFORMMAGMFQVFTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSRSRV 660
Db      601  RAFORMMAGMFQVFTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSRSRV 660
Qy      661  MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSLVLQVVPFLVLVSVGADNIFIFYLE 720
Db      661  MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGIRSLVLQVVPFLVLVSVGADNIFIFYLE 720
Qy      721  YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFFLGALTPMPAVRTPALTSLGLAV 780
Db      721  YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSAICFFFLGALTPMPAVRTPALTSLGLAV 780
Qy      781  ILDFLLQMSAFVALLSLDSKQEASRLDVCCKVPOELPPPGQEGEGLLGFFQKAYAPFL 840
Db      781  ILDFLLQMSAFVALLSLDSKQEASRLDVCCKVPOELPPPGQEGEGLLGFFQKAYAPFL 840
Qy      841  LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQBELALPKDSYLLDYFLFLNRYFEVGPV 900
Db      841  LHWITRGVVLVLLFLALFGVSLYSMCHISVGLDQBELALPKDSYLLDYFLFLNRYFEVGPV 900
Qy      901  YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATERPEQSYLAIPASSWVDDFIDW 960
Db      901  YFVTTLGYNPSSEAGMNAICSSAGCNPFSTQKIQYATERPEQSYLAIPASSWVDDFIDW 960
Qy      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db      961  LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Qy      1021  NIKCPKGLAAAYSTSVNLTSDGQVL-----ASRPMAYH 1053
Db      1021  NIKCPKGLAAAYSTSVNLTSDGQVLDTVAILSPRLEYSGTISAHCNLYLLDLSASRPMAYH 1080
Qy      1054  KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAREVPTTINNVFEQVLTILPEGL 1113
Db      1081  KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAREVPTTINNVFEQVLTILPEGL 1140
Qy      1114  FMLSICLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS 1173
Db      1141  FMLSICLVPTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMALWGISYNAVSLINLVS 1200
Qy      1174  AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db      1201  AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
Qy      1234  LIQIIEFFRLNLTLLGLLHGLVFLPVLISYVGPDPVNPALAEQKRAEAAVAAMVVASCP 1293
Db      1261  LIQIIEFFRLNLTLLGLLHGLVFLPVLISYVGPDPVNPALAEQKRAEAAVAAMVVASCP 1320
Qy      1294  NHPSRVSTADNITYVNHSPGSGIKGAGAINFLPNNGRQF 1332
Db      1321  NHPSRVSTADNITYVNHSPGSGIKGAGAINFLPNNGRQF 1359
```

Search completed: April 7, 2006, 19:32:35
Job time : 576 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2006, 19:23:20 ; Search time 77 Seconds
(without alignments)
2380.515 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGAGISNPLPNNGQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 584368 seqs, 137612332 residues

Total number of hits satisfying chosen parameters: 584368

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_AA_New.*
1: /SIDSS/prodata/1/paa/pct_NEW_COMB.pep.*
2: /SIDSS/prodata/1/paa/US06_NEW_COMB.pep.*
3: /SIDSS/prodata/1/paa/US07_NEW_COMB.pep.*
4: /SIDSS/prodata/1/paa/US08_NEW_COMB.pep.*
5: /SIDSS/prodata/1/paa/US09_NEW_COMB.pep.*
6: /SIDSS/prodata/1/paa/US10_NEW_COMB.pep.*
7: /SIDSS/prodata/1/paa/US11_NEW_COMB.pep.*
8: /SIDSS/prodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6909	100.0	1332	6	US-11-301-094-6
2	6872.5	99.5	1359	6	US-11-270-796-22
3	5407	78.3	1333	6	US-11-270-796-3
4	2402.5	34.8	1278	1	PCT-US06-05584-822
5	2402.5	34.8	1278	6	US-11-191-274A-129
6	2402.5	34.8	1278	6	US-11-191-274A-130
7	2402.5	34.8	1278	7	US-11-385-692-2400
8	2402.5	34.8	1278	7	US-11-385-692-2401
9	1341.5	19.4	1383	6	US-11-301-094-2
10	1046	15.1	1274	6	US-11-301-094-4
11	782.5	11.3	419	6	US-10-953-349-21065
12	588.5	8.5	967	6	US-10-461-673-16739
13	583.5	8.4	1182	6	US-11-332-764-2
14	583.5	8.4	1182	6	US-11-337-244-149
15	482	7.0	891	8	US-60-772-265-1197
16	455	6.6	1358	8	US-60-772-265-241
17	437	6.3	831	7	US-11-360-355-133805
18	349	5.1	783	6	US-10-461-673-16750
19	325	4.7	204	7	US-11-360-355-141351
20	261	3.8	575	7	US-11-360-355-120785
21	252	3.6	505	6	US-10-461-673-16871
22	229.5	3.3	465	7	US-11-360-355-141367
23	225	3.3	542	6	US-11-214-063A-1670
24	222.5	3.2	422	7	US-11-360-355-141365
25	219	3.2	1137	6	US-10-461-673-16889

26	219	3.2	1330	6	US-10-461-673-10541	Sequence 10541, A
27	190	2.8	126	7	US-11-360-355-141352	Sequence 141352, A
28	152.5	2.2	632	6	US-11-293-697-3548	Sequence 3548, Ap
29	141	2.0	164	7	US-11-360-355-127069	Sequence 127069, A
30	139.5	2.0	484	7	US-11-360-355-120169	Sequence 120169, A
31	135.5	2.0	252	7	US-11-360-355-140206	Sequence 140206, A
32	126	1.8	808	6	US-10-724-972B-7050	Sequence 7050, Ap
33	123.5	1.8	171	7	US-11-360-355-149589	Sequence 149589, A
34	121.5	1.8	1043	6	US-10-536-606-20	Sequence 20, Appl
35	120.5	1.7	2426	6	US-11-203-806A-11	Sequence 11, Appl
36	118.5	1.7	768	6	US-11-214-063A-2044	Sequence 2044, Ap
37	118.5	1.7	788	6	US-11-214-063A-1692	Sequence 1692, Ap
38	117.5	1.7	619	8	US-60-732-162-1828	Sequence 1828, Ap
39	116.5	1.7	619	6	US-11-312-958-46	Sequence 46, Appl
40	116.5	1.7	619	6	US-10-461-673-12167	Sequence 12167, A
41	114	1.7	488	8	US-60-752-355-45206	Sequence 45206, A
42	114	1.7	985	6	US-11-293-697-2874	Sequence 2874, Ap
43	112.5	1.6	697	6	US-10-703-799B-226	Sequence 226, App
44	111	1.6	201	7	US-11-360-355-152528	Sequence 152528, A
45	111	1.6	958	8	US-60-781-953-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-11-301-094-6
; Sequence 6, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-094-6

Query Match	100.0%	Score	6909;	DB	6;	Length	1332;
Best Local Similarity	100.0%	Pred. No.	0;				
Matches	1332;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTHIQPGYCAFYDECGKNPELSGSLMTLSNVCLSN	60				
Db	1	MAEAGLRGWLWALLRLAQSEPYTTHIQPGYCAFYDECGKNPELSGSLMTLSNVCLSN	60				
Qy	61	TPARKITGDHLLILQKICPRLYTGNTQACSAQOLVSLASLSITKALLTRCPACSNF	120				
Db	61	TPARKITGDHLLILQKICPRLYTGNTQACSAQOLVSLASLSITKALLTRCPACSNF	120				
Qy	121	VNLCHNTCSPNOSLFINVTRVAQIGAGOLPNAVAYEAFYOHSPAEQSYDSCSRVRPAA	180				
Db	121	VNLCHNTCSPNOSLFINVTRVAQIGAGOLPNAVAYEAFYOHSPAEQSYDSCSRVRPAA	180				
Qy	181	ATLAVGTWCGYVGSALCNQARWLNFGQDGTGNGLAPLDTTFHLLBPQAVGSGIQPLNGV	240				
Db	181	ATLAVGTWCGYVGSALCNQARWLNFGQDGTGNGLAPLDTTFHLLBPQAVGSGIQPLNGV	240				
Qy	241	ARCNESQDDVATCSQDCAASCPAIARQALDSTFYLGQMPGSLVLIILICSVFVVTI	300				
Db	241	ARCNESQDDVATCSQDCAASCPAIARQALDSTFYLGQMPGSLVLIILICSVFVVTI	300				
Qy	301	LLVGRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQPGQGTWASWPLTILVLSV	360				
Db	301	LLVGRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQPGQGTWASWPLTILVLSV	360				

QY 361 IPVVALAAGLVFTLTTDPVLSAPNSQARSEKAFHQHFGPPFRNTQVILTAQNRSY 420
DB 361 IPVVALAAGLVFTLTTDPVLSAPNSQARSEKAFHQHFGPPFRNTQVILTAQNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQVWSPQARNSLQDICYAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQVWSPQARNSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKHLFCYANAPLTFKDGTLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKHLFCYANAPLTFKDGTLAL 540
QY 541 SCWADYGAPVFPFLAIGYKGVSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
DB 541 SCWADYGAPVFPFLAIGYKGVSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFORMMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
DB 601 RAFORMMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATLGLGGVAVVILGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGLLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVPGAPV 900
DB 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFVPGAPV 900
QY 901 YFVTTLGYNFSSEAGNATCSSAGCNFSFTQKIQVATEFPQSYLAIIPASSWDDFDW 960
DB 901 YFVTTLGYNFSSEAGNATCSSAGCNFSFTQKIQVATEFPQSYLAIIPASSWDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNKMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAFEPFPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFEPFPPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
DB 1141 NLLSIWMLVDTVGFMAWDISYNAVSLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATTSMGSAVPAGVAMTNLPGLVGLAKAQLOIIPFRLNLLITLGLHLGVFLPV 1260
DB 1201 AKEATTSMGSAVPAGVAMTNLPGLVGLAKAQLOIIPFRLNLLITLGLHLGVFLPV 1260
QY 1261 ILSYVGPVNPALALQKRAEAAVAVMVASCPNHPSTRYSTADNIIYVNHSPFGSIKGAGA 1320
DB 1261 ILSYVGPVNPALALQKRAEAAVAVMVASCPNHPSTRYSTADNIIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFPNNGRQF 1332
DB 1321 ISNFPNNGRQF 1332

RESULT 2

US-11-270-796-22

; Sequence 22, Application US/11270796

; GENERAL INFORMATION:

; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; TITLE OF INVENTION: RESTORATION OR MIMICRY PF p16 Ink4a ACTIVITY
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-270-796-22

Query Match 99.5%; Score 6872.5; DB 6; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWLLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVNCSLN 60
DB 1 MAEAGLRGWLWLLRLAQSEPYTTIHOPGYCAFYDECGKNPELSGLMTLSNVNCSLN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVSEASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
DB 121 VNLCHNTCSNQSLFINVTRVAOLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAGVTCMGVYGALCNAORWLNFGQDTGNGLAPLDTITFHLLEBQAVGSGIOPLNEGV 240
DB 181 ATLAGVTCMGVYGALCNAORWLNFGQDTGNGLAPLDTITFHLLEBQAVGSGIOPLNEGV 240
QY 241 ARCNEQDQDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNEQDQDVATCSCQDCAAPALPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLIGQPFQGWGTWASWPLTTLVLSV 360
DB 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLIGQPFQGWGTWASWPLTTLVLSV 360
QY 361 IPVVALAAGLVFTLTTDPVLSAPNSQARSEKAFHQHFGPPFRNTQVILTAQNRSY 420
DB 361 IPVVALAAGLVFTLTTDPVLSAPNSQARSEKAFHQHFGPPFRNTQVILTAQNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQVWSPQARNSLQDICYAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQVWSPQARNSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKHLFCYANAPLTFKDGTLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGTSQVMDKHLFCYANAPLTFKDGTLAL 540
QY 541 SCWADYGAPVFPFLAIGYKGVSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
DB 541 SCWADYGAPVFPFLAIGYKGVSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFORMMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
DB 601 RAFORMMAGMFQVTFABRSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATLGLGGVAVVILGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKPQELPPPGQSGLLGFFQKAYAPFL 840

Db	781	ILDELLQMSAFVALLSLDSKQREASRLDCCVCKQDELPPQGGEGLLGFTQKAYAPFL	840
Qy	841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	900
Db	841	LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	900
Qy	901	YFVTTILGYNFSEAGNNAICSSAGCNPFSTQKIQVATEPFPQSXYLAIPASSWVDDFIDW	960
Db	901	YFVTTILGYNFSEAGNNAICSSAGCNPFSTQKIQVATEPFPQSXYLAIPASSWVDDFIDW	960
Qy	961	LTPESSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDLRP	1020
Db	961	LTPESSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDLRP	1020
Qy	1021	NIKCPKGGLLAAYSTSVNLTSDDQVL-----ASRPMAYH	1053
Db	1021	NIKCPKGGLLAAYSTSVNLTSDDQVLDTVAILSPREYSGTISAHNCNLYLLDSASRPMAYH	1080
Qy	1054	KPLKNSQDYTEALRAARELAANITADLRKVPQDPAFVFPYTIITNVFYEQYTLTILPEGL	1113
Db	1081	KPLKNSQDYTEALRAARELAANITADLRKVPQDPAFVFPYTIITNVFYEQYTLTILPEGL	1140
Qy	1114	FMLSCLVPTFAVSCLLIGLDRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLYS	1173
Db	1141	FMLSCLVPTFAVSCLLIGLDRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLYS	1200
Qy	1174	AVGMSVEFVSHITRSPAIKTPWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQ	1233
Db	1201	AVGMSVEFVSHITRSPAIKTPWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQ	1260
Qy	1234	LQIIPFFRLNLTITLGLLHGLVFLPVILSVYVGPVNPALALEQKREABEAAVAVASCP	1293
Db	1261	LQIIPFFRLNLTITLGLLHGLVFLPVILSVYVGPVNPALALEQKREABEAAVAVASCP	1320
Qy	1294	NHPSRVSTADNIVNHSFPGSIGKAGAIINFLPNNGRQF	1332
Db	1321	NHPSRVSTADNIVNHSFPGSIGKAGAIINFLPNNGRQF	1359
RESULT 3			
US-11-270-796-3			
; Sequence 3, Application US/11270796			
; GENERAL INFORMATION:			
; APPLICANT: Dong, Jjianli			
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAP AND			
; TITLE OF INVENTION: RESTORATION OR MIMICRY PP p16 Ink4a ACTIVITY			
; FILE REFERENCE: 02420/1201581-US1			
; CURRENT APPLICATION NUMBER: US/11/270,796			
; PRIOR FILING DATE: 2005-11-08			
; PRIOR APPLICATION NUMBER: 60/592,592			
; PRIOR FILING DATE: 2004-07-30			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: Patentin version 3.3			
; SEQ ID NO 3			
; LENGTH: 1333			
; TYPE: PRT			
; ORGANISM: Mus musculus			
US-11-270-796-3			
Query Match 78.3%; Score 5407; DB 6; Length 1333;			
Best Local Similarity 77.2%; Pred. No. 0;			
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRGHWLLALLRLAQSPYTIHQPGYCAFYDECGKNPELSGSLMTUNSVCSLNTPA	63
Db	3	AAWQGWLLWALLNSAQGBLYTPTHKAGCTFYEECGKNPELSGSLTSLNCSLNTPA	62
Qy	64	RKITGDHLLQKICPRLYTGN-TOACCSAKOLYSLEASLTITKALLTRCPACSDNFVN	122
Db	63	RHVTGDHALLQVCPRLYNGENDTYACCSTKQLVSLDSSLSTITKALLTRCPACSDNFVS	122
Qy	123	LHCHNTCSNQSLFINVTRVAGLQAGQLPAVVAYEAFYQHSFABQSDYSCSRVRPAAAT	182

Db	123	IHCNHTCSPDQSLFINVTRVQDQQLPAVVAYEAFYQHSFABQSDYSCSRVRIPAAAS	182
Qy	183	LAVGTMCGYVGSALCNAORWLNFGQDTGNGLAPLDITPHLLRPGQAVGSGIOPNBSGAR	242
Db	183	LAVGTMCGYVGSALCNAORWLNFGQDTGNGLAPLDITPHLLRPGQAVGSGIOPNBSGAR	242
Qy	243	CNESQGDVATCSCDCAASCPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTILL	302
Db	243	CNESQGDVATCSCDCAASCPAIARPAQALDSTFYLGQMPGSLVLIILCSFVAVVTILL	302
Qy	303	VGFRVAPARDKSNWDPKGTSLSKLSFSTLILGQFQGHGTWVASHPLILVLSVIP	362
Db	303	VYLRVASNRNKTKAGSQEAPNLPRKRFPSPHVLGRFPFESWGTTRVASFVPLTVLASFIV	362
Qy	363	VVALAAGLVFTLTDPVWELWSAPNSQARSSEKAFHDQHPGPFRTNOVLTATPNRSSYRY	422
Db	363	VVALAAGLVFTLTDPVWELWSAPNSQARSSEKAFHDQHPGPFRTNOVLTATPNRSSYRY	422
Qy	423	DSLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLELLEL	482
Db	423	DSLLGPKNFGSILDLLELLELLELLELLELLELLELLELLELLELLELLELLEL	482
Qy	483	YDCCINSILQYQNNRTLLLTANQTLMGQTSQVQWQKHFLYCANAPLTFKQGTALALSC	542
Db	483	YDCCINSILQYQNNRTLLLTANQTLMGQTSQVQWQKHFLYCANAPLTFKQGTALALSC	542
Qy	543	MADYCAPVFPPLAIGSYGKQVSEARALMTPLSNYYPAGDPRLAQAQKLEWBAFLBEMRA	602
Db	543	MADYCAPVFPPLAIGSYGKQVSEARALMTPLSNYYPAGDPRLAQAQKLEWBAFLBEMRA	602
Qy	603	FORMMAGMFQVFTTAAERSLEDEINRTTAAEDLPATSYIVIFLYISLALSGSYSSSRVWV	662
Db	603	FORMMAGMFQVFTTAAERSLEDEINRTTAAEDLPATSYIVIFLYISLALSGSYSSSRVWV	662
Qy	663	DSKATILGCGVAVVILGAVMAAGPFSYLGIRSLVILQVVPFLVSVGADNIPFVLEYQ	722
Db	663	DSKATILGCGVAVVILGAVMAAGPFSYLGIRSLVILQVVPFLVSVGADNIPFVLEYQ	722
Qy	723	RLPREPGRPREVHIGRALGRVAPSMILCSLSAICFFELGALTMPAVRTFALTSGLAIVL	782
Db	723	RLPREPGRPREVHIGRALGRVAPSMILCSLSAICFFELGALTMPAVRTFALTSGLAIVL	782
Qy	783	DELLQMSAFVALLSLDSKQREASRLDCCVCKQDELPPQGGEGLLGFTQKAYAPFLH	842
Db	783	DELLQMSAFVALLSLDSKQREASRLDCCVCKQDELPPQGGEGLLGFTQKAYAPFLH	842
Qy	843	WITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	902
Db	843	WITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	902
Qy	903	VTTILGYNFSEAGNNAICSSAGCNPFSTQKIQVATEPFPQSXYLAIPASSWVDDFIDWLT	962
Db	903	VTTILGYNFSEAGNNAICSSAGCNPFSTQKIQVATEPFPQSXYLAIPASSWVDDFIDWLT	962
Qy	963	P-SSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDLRN	1021
Db	963	P-SSCCRLYISGPNKDKFCPTVNSLNCIKNCMSITMGSVRPSVEQFKYLPWFNDLRN	1021
Qy	1022	IKCPKGGLLAAYSTSVNLTSDDQVLASRFMAYHKLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGGLLAAYSTSVNLTSDDQVLIASQFMAYHKLKNSQDYTEALRAARELAANITADLR	1082
Qy	1082	KVPQDPAFVFPYTIITNVFYEQYTLTILPEGLPMLSLCLVPTFAVSCLLIGLDRSGLLN	1141
Db	1083	KVPQDPAFVFPYTIITNVFYEQYTLTILPEGLPMLSLCLVPTFAVSCLLIGLDRSGLLN	1142
Qy	1142	LLSIWMLVDTVGFMAWLDISYNAVSLINLYSNAVSLINLYSNAVSLINLYSNAVSLINLYS	1201
Db	1143	LLSIWMLVDTVGFMAWLDISYNAVSLINLYSNAVSLINLYSNAVSLINLYSNAVSLINLYS	1202
Qy	1202	KEATISMGSAVPAVAMTNLPGILVLGLAKAQIQLFFPRLNLTITLGLLHGLVFLPVI	1261
Db	1202	KEATISMGSAVPAVAMTNLPGILVLGLAKAQIQLFFPRLNLTITLGLLHGLVFLPVI	1261

Db	1203	KDATIFMGSNVPAGVAMTNFPGLILGPAQALIQIFPFRNLLITLLGLLHGLVFLPVV	1262
Qy	1262	LSYVGDVNPALALOKRABEAAVAVMVASCPNHPRSRSTADNITYVNHSPGSG-IKGAGA	1320
Db	1263	LSYLGDVNPQALVLEBKATEA-AMVSBPSCPQYPPADANTSDYVNGFNPFIPEINA	1321
Qy	1321	ISNFLPNNGROF	1332
Db	1322	ASSSLPKSDQKF	1333
RESULT 4			
PCT-US06-05584-822			
; Sequence 822, Application PC/TUS0605584			
; GENERAL INFORMATION:			
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.			
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION,			
; TITLE OF INVENTION: ASSESSMENT, PREVENTION, AND THERAPY OF CANCER			
; FILE REFERENCE: DFG-064.25(25992-6425)			
; CURRENT APPLICATION NUMBER: PCT/US06/05584			
; CURRENT FILING DATE: 2006-03-02			
; PRIOR APPLICATION NUMBER: 60/690,064			
; PRIOR FILING DATE: 2005-06-13			
; PRIOR APPLICATION NUMBER: 60/654,227			
; PRIOR FILING DATE: 2005-02-17			
; NUMBER OF SEQ ID NOS: 848			
; SOFTWARE: PatentIn version 3.3			
; SEQ ID NO 822			
; LENGTH: 1278			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
PCT-US06-05584-822			
Query Match 34.8%; Score 2402.5; DB 1; Length 1278;			
Best Local Similarity 39.8%; Pred. No. 1.5e-185;			
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;			
Qy	7	RGWLWALLRLRLAQSEPYTHIQPGYCAPYDECG-----KNPELSGSLMTLSNVCSLS	59
Db	4	RLGALGLLLLLCPAQFSQ-----SCVWYGEIGYAKRNCYCEYSG-----	46
Qy	60	NTPAKITGDHLLTLLOKICPLRYTGPTQACCSAKQLVSLSEASLSITKALLTRCPACSDN	119
Db	47	--PPKLPKDGVDLVELCGPFFG-NVSLCCDVRLQTLKONLQLPLQFLSRCPSCFYN	103
Qy	120	FVNLHCHNTCSNPQSLFINVTR----VAQLGAGQLPAVVAEYAFYOHSPAOSYDSCSRV	175
Db	104	LLNLFCBLTCSPRQSQFLNVATATEDYVDPVTNQTKNVKELOYVYVQGSFANAMYNACRDV	163
Qy	176	RVPAATLAVGTWCVGVGSLCNAQRNLNPOGDTGNGLAPLDIT-----PHLLBPQA	228
Db	164	EAPSSNDKALGLLCKGDADA-CNATNWIYMFNKGNGQAPFTITPVFSDFPVH-----	215
Qy	229	VSGIOPLNEGVARCNESQGDVATCSQDCACAPAIARPO-----ALDSTFYLG	279
Db	216	---GMEPMNATKGCDESDEVTAPECSQDCSIVCGPKQPPPPAPWILGLDAMYVM	272
Qy	280	QMPGSLVLIILCSVFAVVWILL-----VGFVAPAPDKSKMVDPKKGTSLS	326
Db	273	WITWMAFLLVFFGAFFAVWCYRKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC	324
Qy	327	DKLSFSTHLLGQFFQGWGTWVASWPLTILVSLVPIPVWALAAGLVFELTTPDVELWSAP	386
Db	325	DPVSAAPFEGCLRLRFTRWGSCFVRNPGCVIFFSLVFIACSSGLVFRVVTNPDVLSAP	384
Qy	387	NSOARSEKAFDQHFQFFRTNOVILTPARNRSSYRYDSLILGPK-NFSGILDLDLLELL	445
Db	385	SSQARLEKEVPDQHFQFFRTEQLIRAPLTDKHIYQYPSGADVPFGPPLDIQLHQVL	444
Qy	446	ELQRLRHLOWSPEAQRNISLQDICYAPLNPDNTSLYDCINSLLQYFQNNRTLLLLTA	505
Db	445	DLQIAIEN--ITASYDNETVTLDQICLAPLSPYNT---NCTILSVLNYFQNSHVLDDHK	499

Qy	506	NQTLMGQTSQVDWKDHFLYCANAPLTPKOGTALALSCMADYDGPVFPFLAIGGYKKDYS	565
Db	500	GDDEF---VYADYTHFLYCVRAPASLNDTSLLDHPCLGTFGGVPFWLVLGGYDNDYN	556
Qy	566	EAEALIMTFSLNNYPAGDPRLAQAOKLWEAFLEEMRAFORMAGMFQVTFTAERSLDEI	625
Db	557	NATALVITFPVNNYNDTEKLQRAQAWEKEFINVKYKN---PNLTISFTAERSIEDEL	613
Qy	626	NRTTAEDLPFATSYIVIFLYISLALSYSSWSVMVDKATLGLGGVAVVLGAVMAAMG	685
Db	614	NRESDDVFTTVVISALMFLYISLALGHISCRLLVDSKVSIGIAGLILVSSVACSLG	673
Qy	686	FFSYLGIRSSILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPPGEPREHVHTGRALGRVAP	745
Db	674	VFSYIGLPLTLVIEVIPFLVAVGVDNIFILVQAYQORDERLQGETLDQQLGRVLGEVAP	733
Qy	746	SMLLCSLSEALCFPLGALTMPAVRTALTSLGLAVILDFLLOMSAFVALLSIDSQROAS	805
Db	734	SMFLSSFSSETVAFFLGALSVMNAVHTFSFLAGLAVFDLLOITCFVSLGLDIKRQBN	793
Qy	806	RLDVCCCVKPOELPPGO-GEGLLGPFOKAYAPFLHWHITRGVVLILFLALFGVSLYSX	864
Db	794	RLDIFCCVGRGAEDGTSSVOASESCLFRFFKNSYSPULLKQWNPPIVIAIFVGVLSFIAVL	853
Qy	865	CHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYFVTTLYGNFSSSEAGNAICSSAG	924
Db	854	NKVDIGLQSLMPDSDVMVDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVCQGMG	913
Qy	925	CNPFSTFKIQIYATEFPQSYLAIPASSWVDDFDWLTP-SSCCLELYISGNKDKPCPST	983
Db	914	CNNDSLVQOIENAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRNV---DNITDQFCNAS	970
Qy	984	VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFNLNDRENIKCPKGLLAAYSTSVN--LTS	1040
Db	971	VDPACVR-CRPLPEGHQRPQGGDFRFLPMFLSDNPNPKCGKGGHAAYSANVILGH	1029
Qy	1041	DQQLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPQDTPAFEVPTITNV	1100
Db	1030	GTRVGATYFMTVHTVLTQTSADFDALKKARLIASNVT-ETWINGS--AYRVFPYSVYV	1086
Qy	1101	FYEQLVTLPEGLFMLSCLVPTFAVSCLLGLDLSGLLNLSTVMTILVDTVGFMAIWD	1160
Db	1087	FYEQLVTLTIDDTIFNLGVSGLAIFLVTVMLGCELWSAVIMCATAMVLNMFVGNWLMG	1146
Qy	1161	ISYNAVSLNLVSAVGMSEVFSHTSRFAISTKTWLERAKEATISMGSAFAGVAMTN	1220
Db	1147	ISLNAVSLVNLVMSGIVSEFCSHITRAFTVSMKGSRVERAEALAHMGSSVFSGITLTK	1206
Qy	1221	LPGILVLGLAKAQLIQIFPFRNLLITLLGLLHGLVFLPVILSYVGPDVNPA	1272
Db	1207	FGIIVVLAFAXSQIFQIFVFRMYLAMVLIGATHGLIFLPLVLLSYIGPSYNKA	1258
RESULT 5			
US-11-191-274A-129			
; Sequence 129, Application US/11191274A			
; GENERAL INFORMATION:			
; APPLICANT: Applera Corporation			
; APPLICANT: Bruno DOMO			
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF			
; FILE REFERENCE: CL001536PROV			
; CURRENT APPLICATION NUMBER: US/11/191,274A			
; CURRENT FILING DATE: 2005-07-28			
; NUMBER OF SEQ ID NOS: 334			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 129			
; LENGTH: 1278			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-11-191-274A-129			
Query Match 34.8%; Score 2402.5; DB 6; Length 1278;			
Best Local Similarity 39.8%; Pred. No. 1.5e-185;			

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLLAQSPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCYAGDKYNCYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGTNTQACCSAKQLVSLBASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGVLDVQELCPGFFG--NWSLCCDVRLQTLKONLQPLQFLSRCPSCFTN 103

QY 120 FVNHLCHNTCSNQSFLINVTNTR-----VAQLGAGQLPAVVAYEAFYQHSFASQSYDSCSRV 175
DB 104 LNLFLCBLTCSPRQSQFLNVTATEDYVDPVTNQTNTNKBELQYVYGQSFANAMYNACRDV 163

QY 176 RVPAAATLAVGTMCYGYGALCNAQRWLNFGQDGTGNGLAPLDIT-----PHILLEPQGA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVFSDPPVH-----215

QY 229 VSGSQPLNEGVARCNESQGDVATCSCDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GHEPMNATKGCDESVDVTPACSCQDCSIVCGPKQPPPPAPMTILGLDAMYIM 272

QY 280 QMPGSLVLIILCSVFVAVVTILL-----VGFVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVFPFGAFFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHTLLQOFFQCGTGWASWPLTILVSLVPPVVALAAGLVFTELTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLFRMGSCFVRNPGCVIPFSLVFIATACSSGLVFRVTTNPVDLWSAP 384

QY 387 NSQARSEKAFDHQGPFRFTNOVILTAPNRSYRYDSLLGLPK-NFSGILDLLELL 445
DB 385 SSQARLEKEYFDHQGPFRTEQLIIRAPLTDKHIYQYPYPSGADVPPGPDLDIQLHQL 444

QY 446 ELQERLRHLQVMSPEAQRNLSLQDICYAPLNDNTSLYDCCINSLLQYFQNNRTLLTLTA 505
DB 445 DIQIALEN--ITASVDNETVTLQDICLAPLSYNT---NCTLSLVNIYFQNSHVDHKK 499

QY 506 NOTLMGQTSQVDKOHFLYCANAPLTFKDGITALALSQADYAGVPFPFLAIGYKGYKS 565
DB 500 GDDEF---VYADYHTHFLYCVRAPASLNDTSLHDPCLTGTFGPPVPLVLGYDDQNYN 556

QY 566 EAEALIMTSLNYPAGDPRLOAKLWBEAFLEEMRAFQORMAGMFWQVTTFAERSLEBI 625
DB 557 NATALVITPPVNNYNDTEKLOQAQAEKEFINFKVNYKN---PNLTISFTAERSIEDBL 613

QY 626 NETTAEDLPFATSYIVILYISLALGSSYSSRWVWDSKATILGGVAVVLGAVMAWG 685
DB 614 NRESQSDVPTVVISYAIMFLYISLALGHYKSCRLLVDSKVSLSLGIAGLILVLSSVACSIG 673

QY 686 PFSYLGISSVLQVPPFLVSVGADNIFIFVLEYQRLPRRPRGPREVHIGRALGRVAP 745
DB 674 VFSYIGLPLTLVIEVIFPLVLAVGVDNIFILVQAYQDERLOGETLDOQLGRVLEVP 733

QY 746 SMLLCSLSBAICFFIGALTPMPARTFALTSGVLVDFLLQMSAFVALLSLSKREQAS 805
DB 734 SMFLSSFSFETVAFFLGALSVPVAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIKREQN 793

QY 806 RLDDVCCVKPQLPPPGQ--GEGILLGPFQKAYAPLHLLWITRGVLLFLFALPGVLSYSM 864
DB 794 RLDFPCCVRGABDGTSSVQASECLFRFFKNYSYSPILLKDDMMRPVIAIFVGLVLSFIAVL 853

QY 865 CHISVLGDELALPKDSYLLDPLFLNRYFEVAGPVYFVTTILGYNFPSSBAGNACSSAG 924
DB 854 NKVDIGLQSLSWPDDSYNDVIFKISQYLHAGPPVYFVLEGHYDTSKGNQWCGMG 913

QY 925 CNMFSTQKIQVATPEPQSYLAIPASSWDDFDIWLTP--SSCCRLYISGPNKDFPCPST 983
DB 914 CNNDLSVQOIFNAALQDNVTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970

QY 984 VNSINCLKNCMSIT--MGSVRSVPEQHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VYDPACVR--CRELTPPEGKORPGQDFMRFLPMLFSLDNPNPKCGKGHAAYSSAVNILLGH 1029

RESULT 6

US-11-191-274A-130
; Sequence 130, Application US/11191274A
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LONG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001536PROV
; CURRENT APPLICATION NUMBER: US/11/191,274A
; CURRENT FILING DATE: 2005-07-28
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 130
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-191-274A-130

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLLAQSPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCYAGDKYNCYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGTNTQACCSAKQLVSLBASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGVLDVQELCPGFFG--NWSLCCDVRLQTLKONLQPLQFLSRCPSCFTN 103

QY 120 FVNHLCHNTCSNQSFLINVTNTR-----VAQLGAGQLPAVVAYEAFYQHSFASQSYDSCSRV 175
DB 104 LNLFLCBLTCSPRQSQFLNVTATEDYVDPVTNQTNTNKBELQYVYGQSFANAMYNACRDV 163

QY 176 RVPAAATLAVGTMCYGYGALCNAQRWLNFGQDGTGNGLAPLDIT-----PHILLEPQGA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVFSDPPVH-----215

QY 229 VSGSQPLNEGVARCNESQGDVATCSCDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GHEPMNATKGCDESVDVTPACSCQDCSIVCGPKQPPPPAPMTILGLDAMYIM 272

QY 280 QMPGSLVLIILCSVFVAVVTILL-----VGFVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVFPFGAFFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHTLLQOFFQCGTGWASWPLTILVSLVPPVVALAAGLVFTELTDPVELWSAP 386
DB 325 DPVSAAFEGCLRLFRMGSCFVRNPGCVIPFSLVFIATACSSGLVFRVTTNPVDLWSAP 384

QY 387 NSQARSEKAFDHQGPFRFTNOVILTAPNRSYRYDSLLGLPK-NFSGILDLLELL 445
DB 385 SSQARLEKEYFDHQGPFRTEQLIIRAPLTDKHIYQYPYPSGADVPPGPDLDIQLHQL 444

QY 446 ELQERLRHLQVMSPEAQRNLSLQDICYAPLNDNTSLYDCCINSLLQYFQNNRTLLTLTA 505

```
Db 445 DLQTAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLNFQNSHVLDDHKK 499
Qy 506 NQTLMGOTSQVDWKDHLFLYCANAPLTFKDGFTALALSNCADYGAVPFPFLAIGGKYGKDY 565
Db 500 GDDPF--VYADYHTFLYCYRAPASLNDTSLHDPCLTGTPGVPFPLVGLGGYDDQNYN 556
Qy 566 EBAALIMFTSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFOVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFINFKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSIVIFLYISLALSGSYSSWSRVMVDSKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSDFVTWVISYAINFLYISLALGHIKSCRRLLVDSKVSLSIAGILVLSSVACSIG 673
Qy 686 PFSYIGIRSSILVLOQVFPFLVLSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VPSYIGLPLTLVIEVIFPLVLAVGVDNIFILVQAYQDERLQGETLQOGLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFLGALTPMPAVTFEALTSGLAVILOFLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSFSFSETVAFFLGALSVMNAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQEK 793
Qy 806 RLDVCCVKPQBLPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASECLFRFFKNSYSPLLDKDWMRPVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYLFLFLNRYFEVGAPVYFVTTILGNFSPSEAGNNAICSSAG 924
Db 854 NKVDIGLQDLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEBGHDYTSKGNMVCGGMG 913
Qy 925 CNNFSTQKIQVATPEPQSYLAIIPASSWDDFIDWLTP--SSCCRLYISGPNKDFCPS 983
Db 914 CNNDLSVQOIFNAALDNTYRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR--CRPLTPBEGKRPQGGDFMRFLPFLSNDPNKPCGKGHAAYSAMVILGH 1029
Qy 1041 DQOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPDGTDPAFEPFYITNV 1100
Db 1030 GTFVGATYFWYHTVTLQTSADFDALKKARLIASNT--ETWINGS--AYRVEPYSVTV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLPPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFALWD 1160
Db 1087 FVEQYLTIIDTIFNLGVSIGAIFLVTWLLGCELMSAVIMCATIAMVLVNMFGVMWLG 1146
Qy 1161 ISUNAVSLNLVSAGMSVEFVSHITRSPAISTKPTMLERAKEATISMGSAVPAVAMTN 1220
Db 1147 ISUNAVSLNLVMSCGISVEFCSHITRATVSMKGRVEREALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIFFRMLMLTLGLHLGLVFLPVILSYGVPDYNPA 1272
Db 1207 FGGIVVLAFAKSQIFQIFYFRMYLAMVLLGATHGLIFLPLVLLSYIGPSYNKA 1258
```

RESULT 7

```
US-11-385-692-2400
; Sequence 2400, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385,692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2400
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2400
```

```
Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGMILLMALLRLAQSEPYTTIHQPGYCAFYDECG-----KNPELSGLMTLNVSCLS 59
Db 4 RGLAUGLGLLLCLCPAQVFSQ-----SCVMYGECCGAIYDKRYNCEYSG----- 46
Qy 60 NTPARKITGTHLILLOKICRPLTYGPNTOACCSAKQLVSLSEASLITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLCCDVRQOTLKDNILQQLPQFUSRCPCSFYN 103
Qy 120 FVNLHCHNTCSNOSLEFINVTR-----VAQLGAGQLPAVAYEAFYQHSFABQSDSCSRV 175
Db 104 LLMFLCELTSFROSQFLNVTATEDYDVPVNTQKTNKELQYVVGQSFANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCGVYGSALCNAQRLNFGDGTGNGLAPLDIT-----PHLLEPQQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNMIYEMFNKDNQGAQPTTITPVFSDFFVH----- 215
Qy 229 VSGIQLPNEGVARCNESQGDVVATCSQDCAASCAPAIARPQ-----ALDSTFYL 279
Db 216 ---GHEPMNNATKGDSDSDVETAPSCQDCSIVCGPKPQPPPPAPWTIILGLDAMYIM 272
Qy 280 QMPGSLVLIILCSVFAVVTILL-----VGFVRVAPARDKSMVDPKKGTSL 326
Db 273 WITYMAFLVFFGAFFAVWCYRKYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLGLQFQGGTWSWAPLTLVLVSVVVALAAGLVFTLTTDTPVELWAP 386
Db 325 DPVSAAPFGCLRLRLFTRWGSFCVRNPGCVIEFSLVFTITACSSGLVFRVVTNPNVDLWAP 384
Qy 387 NSQARSEKAFHDQHGPFERNQVLTAPNRSSVRYDSLLGLPK--NFGSILLDLDLLELL 445
Db 385 SSQARLEKEYFDQHGPFERTEQLIIRAPLTDKHIYQYPGADVPFGPPDIDILHQLV 444
Qy 446 ELQERLRHLQVMSPEAQARNISLQDICVAPLNPNTSLYDCCINSILLOFQNNRTLLTLTA 505
Db 445 DLQTAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTILSVLNFQNSHVLDDHKK 499
Qy 506 NQTLMGOTSQVDWKDHLFLYCANAPLTFKDGFTALALSNCADYGAVPFPFLAIGGKYGKDY 565
Db 500 GDDPF--VYADYHTFLYCYRAPASLNDTSLHDPCLTGTPGVPFPLVGLGGYDDQNYN 556
Qy 566 EBAALIMFTSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFOVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFINFKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSIVIFLYISLALSGSYSSWSRVMVDSKATLGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSDFVTWVISYAINFLYISLALGHIKSCRRLLVDSKVSLSIAGILVLSSVACSIG 673
Qy 686 PFSYIGIRSSILVLOQVFPFLVLSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VPSYIGLPLTLVIEVIFPLVLAVGVDNIFILVQAYQDERLQGETLQOGLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFLGALTPMPAVTFEALTSGLAVILOFLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSFSFSETVAFFLGALSVMNAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQEK 793
Qy 806 RLDVCCVKPQBLPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASECLFRFFKNSYSPLLDKDWMRPVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYLFLFLNRYFEVGAPVYFVTTILGNFSPSEAGNNAICSSAG 924
Db 854 NKVDIGLQDLSMPDDSYMVDYFKSISQYLHAGPPVYFVLEBGHDYTSKGNMVCGGMG 913
Qy 925 CNNFSTQKIQVATPEPQSYLAIIPASSWDDFIDWLTP--SSCCRLYISGPNKDFCPS 983
Db 914 CNNDLSVQOIFNAALDNTYRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
```



```
Db 971 VDPACVR-CRPLTPGKORPOGGDFMFLPMLSDNPNPKCGKGGHAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMYTHVLTQTSADFALKKARLIASNT-ETMGINGS--AYRVPYSVFV 1086
Qy 1101 FYEQYLTILPEGLFMLSLCLVPTFAVSCULLGLDLRSLNLLSIYMLIVDTVGFMAWD 1160
Db 1087 FYEQYLTIIIDDTIFNLGSLGAIPLVTMVLGCELMSAVIMCATIAMVLVNMFGVMWLG 1146
Qy 1161 ISYNAVSLNLVSAVGMSEVFSHTSRFAISKTPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKSRVERAEALAHMGSSVPSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRMLNLLITLGLHLGLVPLPVILSYVGPVNPA 1272
Db 1207 FGGIVVLAPAKSQIQIFFRMYLANVLLGATHGLIFLPLVLSYIGPSVKA 1258

RESULT 8
US-11-385-692-2401
; Sequence 2401, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USBS THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385,692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2401
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2401

Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLRAQSEPYTHIQPGYCAPYDRGG-----KNPELSGSLMTLSNVCSLUS 59
Db 4 RGLALGLLLLLCPAQVFSQ-----SCWYGEGBIAYGDKRYNCBYSG----- 46

Qy 60 NTPARKITGDHLILQKICPRELYTGNTQACCSAKQLVLEASLSTTKALLTRCPACSDN 119
Db 47 ---PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKONLQLPLQFLSRCPSCFYN 103

Qy 120 FYNLHCHNTCSNQSLFINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCFLTCSPRSQFLNVATEDYDVPVTQTKNVKELQYVYVQSFANAWYACRDV 163

Qy 176 RYPAATLAVGTWCGYVYGALCNAQRWLNFGQDGTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCCGKDADA-CNATNWIEMFNKONGQAPFTITPVFSDFPVH----- 215

Qy 229 VSGSLOPLNEGVARNESQGDVATCSODCAASCPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNNAKGDQESVDEVTAPCSQDCSIVCGPKPQPPPPAPMTILGLDAMYIM 272

Qy 280 QMPGSLVLIILILCSFVAVVTILL-----VGRFVAPARDKSKWDPKKGTSLS 326
Db 273 WITYMAFLVFFGAPFAVWCYKRYFVSEYTPIDNSIAFSV-NASDKGE-----ASCC 324

Qy 327 DKLSFSTHTLLGQFGWGTWASVWMLTTLVLVSVIPVVALAAGLVFTLTDPVBLWSAP 386
Db 325 DPVSAFEGCLRLFRWGSFCVRNPGCVIFPSLVFITAACSSGLVFRVTRTNPVLDWSAP 384

Qy 387 NSQARSEKAFHDOHGPFRTHQVILTAPNRSYYSYDLSLLGPK-NFSGILDLDLLELL 445
Db 385 SSQARLEKEYPQHFGPFRTEQLIRAPLTDKHIYQPYPSGADVPFPGPLDIQLHQLV 444
```

```
Qy 446 ELQERLRHLQVWSPEAQRNLSIQDICYAPINPDNTSLYDCINCINSLLOQYFQNNRLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTIQLDICALPLSPNT---NCTILSVNLVYFQNSHVDLHKK 499
Qy 506 NOTLWQOTSQVDKWHFLYCANAPLTFKDGTAALASCMADYGAAPVPPFLAIGYKQKQYS 565
Db 500 GDDFF---YVADYHTHFLYCVRAPASLNDTSLLDHDFCLGTFPGGVPVFWLVGLGDDQNYN 556
Qy 566 EBAELIMTFSLANYPAGDPRLAQAKLWBEAFLEEMRAFORMMAGMFQVFTTFAERSLEDEI 625
Db 557 NATALVIFPVNNYNDTEKQRAQAWKEKFINFKVKNYK---PNLITISFTAERSLEDEL 613
Qy 626 NRTTAEIDLPIFATSYIVIFLYISLALGSYSSRSRVMSDKATLGLGGVAVLGVANMAMG 685
Db 614 NRESDSDFVTWISYAINPLFYISLALGHKSKRLLVDSKVSIGIAGILVLSVACSLG 673
Qy 686 PFESYLGIRSSLVILQVWPELVLSVGADNIFIFVLEFQRLPRRPGPREVHIGRALGRVP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQRDERLQGETLDQOLGRVLGEVAP 733
Qy 746 SMLLCSLSEACIFFLGALTPEMPAVRTFALTGSLAVILDPLLOMSAFVALLSDSKREBAS 805
Db 734 SMFLSPSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQLTFCVSLGLDIDRKQKKN 793
Qy 806 RLDVCCCKVPQBLPPPGQ-GEGLILGFFQKAYAPFLLHWTIRGWVLLLLFLALFGVLSYM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFPKNSVSPILLKXDMRPVIAIFVGLVLSIAVL 853
Qy 865 CHISVGLDQELAPKDSYLLDYLFLNRYFVGAPVYFVTTLGYNPFSSBAGNNAICSSAG 924
Db 854 NKVDIGLDQSLSMPSMDSYMDVYFKSISQYLHAGPPVYFVLEBHDYTSKQGNMVCGGMG 913
Qy 925 CNNEPFTOKIOVATEFPBQSYLAIPASSWVDDFIDWLT-SSCCRLYISGPNKDCPST 983
Db 914 CNDSLVQOIFMAAQLDNTRIGPAPSPMIDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT-MGSRVPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPGKORPOGGDFMFLPMLSDNPNPKCGKGGHAYSSAVNILLGH 1029
Qy 1041 DGQVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATYFMYTHVLTQTSADFALKKARLIASNT-ETMGINGS--AYRVPYSVFV 1086
Qy 1101 FYEQYLTILPEGLFMLSLCLVPTFAVSCULLGLDLRSLNLLSIYMLIVDTVGFMAWD 1160
Db 1087 FYEQYLTIIIDDTIFNLGSLGAIPLVTMVLGCELMSAVIMCATIAMVLVNMFGVMWLG 1146
Qy 1161 ISYNAVSLNLVSAVGMSEVFSHTSRFAISKTPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLVMSCGISVEFCSHITRAFTVSMKSRVERAEALAHMGSSVPSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRMLNLLITLGLHLGLVPLPVILSYVGPVNPA 1272
Db 1207 FGGIVVLAPAKSQIQIFFRMYLANVLLGATHGLIFLPLVLSYIGPSVKA 1258

RESULT 9
US-11-301-094-2
; Sequence 2, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Smith, Mareha
; APPLICANT: Levitan, Diane J
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 1383
```

```
; TYPE: PRF
; ORGANISM: Caenorhabditis elegans
US-11-301-094-2

Query Match      19.4%; Score 1341.5; DB 6; Length 1383;
Best Local Similarity 27.5%; Pred. No. 2e-99;
Matches 369; Conservative 257; Mismatches 561; Indels 155; Gaps 40;

QY 76 KICPLYTGPNQAACSAKQVLSBASISYKALLTRCPACSDNPNVHLCHNTCSNQSL 135
Db 63 EFCPHLLTGDN-KLCTPSQABGLTKQAQARHILGRCPDFNPAKLWCBEFTCSNQD 121
QY 136 FINTRVAQL--GAGOLPAVAYEAF-----YQHS--FAEQSYDSCSVRVPAATAAVG 186
Db 122 FVSIEMKPIEKKEGFTPEYQAEAYVNTVEYRLSTDFAEGFSSCKDVTGGQPALRY- 180
QY 187 TWCYGVGSALCAQWLNFGQDGTGNL-APLDITFHLLPEGQAVSGIQP-LNEGVARCN 244
Db 181 -NC---TSTPCTLTWLEFIGTQNLDLNPIHTKFLLYDPIKTPPSDRSTYMNVTGCD 236
QY 245 ESQDDVATCSQDCQCAACPAIARQALDSTFYLQM-----PGSLVLIILCSFVAVVT 299
Db 237 KSAVGVWPCASTSEC--NKEEYANLIDLDGKTSQTCNVHGIACLNIFVMLAFIGSLAV 294
QY 300 ILLVGF-----RVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFF 341
Db 295 LLCVGFVFTSYDEYDYNLRQTQSGEESPKRNIK-----RTGAWI-----HNFENNA 342
QY 342 QCGTGWASWPLTILVLSVIPVALAAGLVFTTELITDPVELWSAPNSQARSEKAFDHF 401
Db 343 RIGWMAGRNPKSHFFPGICAVLIFCLPGMIYHKESTNVVDMSSPRRARQEWVFNAF 402
QY 402 GPFFRTNQVILTAENRNSYRYDGLLGPKNFGILDLLLELLELLEQLRLHLQWSPBA 461
Db 403 GRPQRYQIMLL--SHRDQSSGKLYP-----VFHKDIFEEFLDILNAIKNISTQSDG 455
QY 462 QNLSLQDLCYAPLNDNTSLVDCNSILQYFQNNRTLLLTANQT-----L 509
Db 456 -RTITLDDVCYRPMGPG---YDCLIMSPNTYFQGNKEHLDMKSNKEETVSEDDADPYF 510
QY 510 MGQTSQVMDKDHLYCANAPLTFKQGTALALSQADYAGAPVFPFLAIGGYKGYSEABA 569
Db 511 SSEATTDENWNNHAAICIDQPMQK--TKSGLSCMGTYGGPSAPNM-VFGKNSTNHQAAS 567
QY 570 LIMTFSLNYPAGDPLAQAQLWEAFLEEMRAFQRRMAGMFQVTFATERSLEDEINRT 629
Db 568 IMMILVQTQ--RTEPBIQAELEWEKFLKFEKREKSPKI--FSPMAERSITDEIENDA 624
QY 630 AEDLPFATSYIVIFLYISALGSY-----SSWSRVMDSKATLGLGVAVVLGVAAM 684
Db 625 KDBIVTVVIALAFILIGVTFSLGRYFVFNQNLWS--ILVHSRICGLMLSVIINLLSFC 583
QY 685 GFPSYLGIRSSVILQVVPFLVLSVGNADNIFIFVLEY--QRLP---RRFCEPREVHIGRA 739
Db 684 GIESMGHHPKVALVQFPVTVLLGVCRTFMVVKYAAQORVSMYSPDQCPEI-VGMV 742
QY 740 LGRVAPSMILCSSEACFPFLGALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSLS 799
Db 743 MAGTMPAMFESSLGCAPSPFICGFTDLPAIRTFCLVAGLAVLDVVVHLCTIFLAFVWD 802
QY 800 KRQASRLDVCCCKPQELPPQGBGL-----LLG-----FFQKAYAPFLLHWI 844
Db 803 QRELNG-----KP-EFPFPYQIKDLLGAYLIGQRATDTFTWTFQFHQVAPFLHHRM 853
QY 845 TRGVVLLFLALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGAFFVT 904
Db 854 TRITGIIFIASFTITVILSSKISVGPQDSMAFTEKSYISTHRYLDKFPDVGPPVFFTV 913
QY 905 TLGVNFSEAGMNAICSSAGCNFSTQKTYATEPPEQSYLAIPASSWDDFIDMLT-P 963
Db 914 DGLDWHRPDPQNKFCFTFPGCSDTSTFGNIMNAYVGHTEQTYLSGEMYNIDNLEISR 973
QY 964 SSCRLYISGPNKDKFCSTVNSL-----NCLKNCM-----SITMSV---RPSVEQFH 1009
```

```
Db 974 SPCKVYVHDN--TFCSTNRNKSALDDKACRTCMDFDYVANSYPKSSIMYHRPSIEVFY 1031
QY 1010 KYLPWFPLNDPNIKCPKGGLAAYSTSVNLTSDGOVLASRFMAVHKPL--KNSQDYTLBALR 1067
Db 1032 RHLRHFLEDTPENGSECVFGRASFKDAISFTSRGRIQASQFMTHKKLSISNSSDFIKAMD 1091
QY 1068 AARELAANITADLRKVPGTDPAPFEVFPYTTITNVFEQYLITLPEGLFMLSCLCYPTFAVS 1127
Db 1092 TARWVSRLERSI-----DDTAHVFAYSKIFPPEQYSTIMPILITTLQTFITVVGVFGII 1145
QY 1128 CLLGLDLRGLGLSIVMLVDTVGFMAWDISYNAVSLINLVSAVGSVFEVSHITR 1187
Db 1146 CVTLGIDVKGAAACAVICQVSNYPHIVAFMVFIFNIPVNAISATNIVMSSGILIEFSVNLK 1205
QY 1188 SPALSTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIOIPEFFRLNLIT 1247
Db 1206 GYACSLQRAKRAESTVGSIGPIILSGPVVVMAGSTWFLSGAHLQIITVTFPKLFLITI 1265
QY 1248 LLGLLHGLVFLPVILSVYVGPVDPNPALALEQKRAEAAVAAVMVASCPNHPSPRVSTADNIYV 1307
Db 1266 VSSAVHALIILPILLAFGGSRGHSSETSTNDNDQHDACVLS--PTAESHISNVEEGIL 1323
QY 1308 N-----HSFEGSIKGAIGAI 1321
Db 1324 NRPSLLDASHILDPLLKAEGGI 1345

RESULT 10
US-11-301-094-4
; Sequence 4, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 4
; LENGTH: 1274
; TYPE: PRF
; ORGANISM: Caenorhabditis elegans
US-11-301-094-4

Query Match      15.1%; Score 1046; DB 6; Length 1274;
Best Local Similarity 25.7%; Pred. No. 1.6e-75;
Matches 345; Conservative 241; Mismatches 587; Indels 168; Gaps 46;

QY 10 LLWALLRLIAQSEPTTHIQPCYCAFYDECCKNPFLSGSLMTLSNVSCLSNTPARKITGD 69
Db 14 VLFLLLIHLALCQ-----AKCYM--TECDGEEDSNHPCKTNKSTYLPITVTRSLNPT 64
QY 70 HLILQLKCPRLTYTGN--TOACCSAKOLVLSBASISITKALLTRCPACSDNPNVHLCHNT 128
Db 65 YNARPEKYSYLVQBEDKAQVCCBELQLKGMTDRISNAATILGSCSPCFDNFAKLWCQFT 124
QY 129 CSPNQSLFNTVRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSVRVVPAATAAVGTM 188
Db 125 CSPDQSKFMKVMET---TGPKNVVVVQMEFKVNRDFVEGLYESCRHTWPFANGIALRLMSL 180
QY 189 CGVYGSALCAQWLNFGQDGTG-NGLA---PLDITFHLLPEGQAVSGIQPLNEGVARCN 244
Db 181 GGVKS-----FENFYGMGTGNLAQSIPIINTEBFQFSRMKNAMNIPITPP-----CH 225
QY 245 ESQDDVATCSQDCQ---AASCPAIARQALDSTFYLQMPGSLVLIILCSFVAVVTIL 301
Db 226 KSAVGVWPCAGDAICDCTNAHQVLDISKVEHLGTQVPHPHFP--DFEWLLKICGCLA-LTVL 283
QY 302 LVGF-----RVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQG-----WGTWVAS 350
```

```
Db 284 LWFILKYSCHRSPNGEGCVYDLGKN-----LEVFEGLCARYANAVTK 330
Qy 351 WPLTILVLSPVVALAAG-LVFTELTTPDELWLSAPNSQARSEKAF-HDQHFGPPFRTN 408
Db 331 HPLIFVSLGLIVAAACCSGNFKFSLTHSDVQVSAADGETRNEKKFIHS--FGPNHRIE 388
Qy 409 QVILTPNRSRYSDSLGPKNFSGILDLDLLELLEQLERLHUVQVSPRAQRNISIQ 468
Db 389 QIFINLPPTT-----KSMFNPLFEEMFQVLVGNIGNLT--ACYGNSSVKLD 432
Qy 469 DICYPALNDNTSLYDCCINSLLQYFQNNRTLLLTANQTLNLCOTSDQVMDKDFLYCANA 528
Db 433 DICYPIDIGN-----HGCALMSPNTYFQNKWTFENAGPPTIDDEIFDDQHWELKYCIRN 488
Qy 529 PLTFKDGDTALALSCMADYGAPVPPFLAIGG-----YKGD-YSEAEALIMTFSLNPNYPAGD 583
Db 489 PLTV--STYSEMSCGFGSPDPLVFGSNESIKGAEMYTYTARTIMITVIRG-----541
Qy 584 PRLAOKLWEAPLEMRAPFORMMAGFOVTTFAERSLEDEINRTAED--LPIFATSYI 641
Db 542 PE-DQAIAMETAFPLNMSRYEMKHANF---TPWTETSAEETHAVETDKIVSVIACAIV 597
Qy 642 VFLYISIALGSV--SSSRVMDVSKATLGLGVAVVLGAVMAAGFFSYLGISSLVIL 699
Db 598 LIWVITMGINHPSSILSALVHHKLLISISAVMISVISVWCISGMFSLFGVHATDNAI 657
Qy 700 QVVPFLVLSVGADNIPFVLEYQRLPRRPGEP-----REVH--IGRALGRVAPSMILCSLS 753
Db 658 VVLPFVITCLGINRIFVIRITQANGCHCYGLPNISYRENNHRIISVNRRSIPIVLNLSI 717
Qy 754 EACFPL-GALTP-----MPAVRTFALTSGLAVILDFLQMSAFVALSLDSKQ-----802
Db 718 CSTCLFLAGGLVFPYVSVPFAVEVFARHAGLAILMDTAFYLLVMLPLFQYDARREMSGRK 777
Qy 803 -----EASRLDVCCVKQELPPQCGEGLLLGFPQKAYAPFLHWTIRGVVLLP 853
Db 778 ETWPWYELNKGKINCLMEAVDGNLRSP-----VDWFKLAIAPLLKIKRIWIATPF 830
Qy 854 LALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFVGAPVYVVTTLGYNFSSE 913
Db 831 FVSLIACYCTCLFPGFNQWAFSETSYLTKHFQNMENINIGPLWFFVSGDVKNHDP 890
Qy 914 AGMAICSSAGCNFNFTQKIQ---YATFPBQSYLAIIPASSWVDDFDIDLTP--SSCRL 969
Db 891 KMQNKFTLAGCDDNSMGNKIRSLAENY-KGNYLHGDVNIWLDLYLQFMHPRGSCCKM 949
Qy 970 YISGPNKDFC-RSTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFLNDRNPKCPKGG 1028
Db 950 -----DGKQFCDFPS--NATHC-SSCSSVASLTTTEYEFYRNLHHFLETPPSIQCAHGG 1001
Qy 1029 LAAYSTSVNLTSDGOVLASRFMAYHKP--LKNSQDYTEALRAARELANITADLRKVPCT 1086
Db 1002 MALAKPAINLNRNGIKQSYFTFPKMLNSISQIYDNRFAKYLADDIEREL-EIPGV 1060
Qy 1087 DPAFEPFPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1146
Db 1061 ----KVYVYSTFPFYEQYLITLSTTVTLVLVLFVAFVTIISLFLRVNLAGSLVTVFVLL 1116
Qy 1147 MLVDTVGFMALWDISYNAVSLINLVSAVGMSEVFSHTSRFAISTKPTWLERAKBATI 1206
Db 1117 SSYLHLMEMMYLIGITVNVVSVINMAMSLGIAVEFFGQMLHGFFYNSKKPKREERAPALV 1176
Qy 1207 SMGSVAFVAGVAMTNLPGLIIV----LGLAKAQILQIFPFRNLALITLGLHLGLVPLV 1262
Db 1177 SNGSTLSGI-----FPAINITAGCLSPADSRLVITFCNQLVGLVCAVHGVVYMTLL 1232
Qy 1263 SYVGPVNPALAEQKRAEEA 1283
Db 1233 AIFGSDFPYQVNSSEESTDEA 1253
```

RESULT 11

```
US-10-953-349-21065
; Sequence 21065, Application US/10953349
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THERY
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953,349
; NUMBER OF SEQ ID NOS: 2004-09-30
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21065
; LENGTH: 419
; TYPE: PR
; ORGANISM: Glycine max
; US-10-953-349-21065

Query Match 11.3%; Score 782.5; DB 6; Length 419;
Best Local Similarity 38.4%; Pred. No. 7.1e-55;
Matches 164; Conservative 88; Mismatches 138; Indels 37; Gaps 10;

Qy 863 SMC-HISVGLDQELALPKDSYLLDYFLNRYFVGAPVYVVTTLGYNFSSEA-GMAIIC 920
Db 2 ALCTRIEAGLEQOIALPRDSYLGQYFSNISEYLRVGPPLYFVK-DYNSLESKHTNQLC 60
Qy 921 SSAGCNFNFTQKIQYATEFPBQSYLAIIPASSWVDDFDIDLTPS--SCRLYISGPNKDK 978
Db 61 SISHCDSNLSLNEISRASLVPTSSYIAKPAASWLDLFWISPEAFSCCKP----TND 116
Qy 979 FCPSTVNSLNCNKMSITMGSV-----RPSVEQFHKYLPWFLNDRPNI 1022
Db 117 YCPDDQPPCCLPDEGCGLGVCCKDCTTCFRHSDLVNDRPSTAQFRELKPLWFLDALPSA 176
Qy 1023 CKPKGGLAAYSTSVNLT--SDGOVLASRFMAYHKPNSQDYTEALRAARELANITADL 1080
Db 177 DCAKGHGAATNSVDLNGYEGVITQASEFRYTHLNRQDVTNARAARDFAISLSS 236
Qy 1081 RKVPTDPAFEPFPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 237 K-----MDIFPYSVFPFPEQYLDIWKALINITVALGAIFV-CLITSSVWSSAI 287
Qy 1141 NLLSVMLVDTVGFMALWDISYNAVSLINLVSAVGMSEVFSHTSRFAISTKPTWLER 1200
Db 288 LLLVLIMILDLGMVMAILQILNNAVSVNLMISGIAVEFCVHIVHAFMWSIGDR-SQR 346
Qy 1201 AKEATISMGSAVPAGVAMTNLPGLIIVGLAKAQILQIFPFRNLALITLGLHLGLVPLV 1260
Db 347 AKTALCTMGASVFGITLTKLVGLVLCFSTSEIFVYVYFQMYLALVIIGLHLGLVPLV 406
Qy 1261 ILSYVGP 1267
Db 407 VLSLFGP 413

RESULT 12
US-10-461-673-16739
; Sequence 16739, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Kyle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
```

```
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunging
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 16739
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-461-673-16739

Query Match      8.5%; Score 588.5; DB 6; Length 967;
Best Local Similarity 21.3%; Pred. No. 1.4e-38;
Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;

QY 334 HT-----LLGQFPQGGTWWASWPLTILVSVIPVVALAAGLVF--TELTDPVELWSAP 386
DB 115 HTDCEGLLSRTFQWLGWQGAHPWIFLLAPLMTAALGTGFLYLPKDEEEDLEEHTYEV 174
QY 387 NSOARSEKAFHDOHFGPFRTNQVILTAPNRSS-----YRYDSLILGPKNFGIL 436
DB 175 GSPAKAERRVQGH-----FTNDSYRFSASRRSTEANFVSLLVVSYSDSLDDPATFAVS 230
QY 437 DLDLLELELELQERLRLHQLQWSPEAQRNISLQDIC--YAPLPNDNTSLYDCCT--NSLLQ 492
DB 231 KLDGAVQDLRV-----AREKGSQIQVQVCARYAL-----CVPPNPILY 270
QY 493 YFQNNRTLLLTANQTLMGQTSQVDWKDHFYCANAPLTFKGTALALSCMADYGA-PVF 551
DB 271 AWQVNTKTLN-----SSISFPAY-----NHGRHPLY 296
QY 552 PFLAIGY-----KGKDYSEAEALIMTFLSNYPAGDPRL-AQAKLMEAEAELEEMRAF 603
DB 297 LTGFFGGYILGSLGWLGMQQLLRKAMKLLYLK---TEDEYDVQSKQWMLTHLDDFTNI 353
QY 604 QRMA-----GMFQVFTFAERSLEDEINRTTAEDLPFATPSYVIFVYISL 649
DB 354 KNILAKKIEVPGVGVLQGGQEKVHFTSLRQLEPEATSVTVIPVPHLAYILILFAVT 413
QY 650 ALGYSYSSWRVWDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQVVPFLVSV 709
DB 414 SCRFD-----CIRNKMCAAFGVSFAFLAWSGFGLLHIGV-PFVIIIVANSFPFLLGV 467
QY 710 GADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAPV 769

; APPLICANT: Wlozetkey, Robert G.
; TITLE OF INVENTION: PATCHED HOMOLOG 2 (PTCH-2) DISRUPTIONS, COMPOSITIONS AND METHODS
; FILE REFERENCE: R1673 CIP/75658.064000
; CURRENT APPLICATION NUMBER: US/11/332,764
; CURRENT FILING DATE: 2006-01-13
; PRIOR APPLICATION NUMBER: US 60/413,543
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 10/669,143
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
; US-11-332-764-2

Query Match      8.4%; Score 583.5; DB 6; Length 1182;
Best Local Similarity 24.2%; Pred. No. 4.7e-38;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SFSTHTLLGQ-----FFQG-----WGTWASWPLTILVSVIPVVALAAGLVTELT 376
DB 20 SSAPHILAGSLQAPLMLRAYFQGLLPSLGRICKGKGVFLGLVAFGALGLRVAVIE 79
QY 377 TDPVELWSAPNSQARSEKAFHQHFG-PFRTNQVILTAPNRSSYRYDSILLGPKNFGSI 435
DB 80 TDLEQLWVEVGRVSQELHYTKELGEEAAYTSQMLI-----QTAHQEGGNVLTPE---A 131

; Sequence 2, Application US/11332764
; GENERAL INFORMATION:
; APPLICANT: Wlozetkey, Robert G.
; TITLE OF INVENTION: PATCHED HOMOLOG 2 (PTCH-2) DISRUPTIONS, COMPOSITIONS AND METHODS
; FILE REFERENCE: R1673 CIP/75658.064000
; CURRENT APPLICATION NUMBER: US/11/332,764
; CURRENT FILING DATE: 2006-01-13
; PRIOR APPLICATION NUMBER: US 60/413,543
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 10/669,143
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
; US-11-332-764-2
```

```
QY 436 LDLDLLELLELQERLHLOVNSPEAQRNISIQIDICY-----APLNPDN-----TSLYDCC 486
D 132 LD-----LHLQAALTASKQVSVYKSGWDLNKKCYKSGVPLIENGMIERMIKLFPCV 184
QY 487 INSLLOYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---PKDGTALA 539
D 185 ILTPLDCFWEGAK---LOGGSAYLPGRPDIOQNTLNDPQOLLEELGPFASLEGFRELDDKA 241
QY 540 LSCMADYGAPVF-----PFLA---IGYKKGKDYs----- 565
D 242 QVGQAYVGRPCLDPPDPPHCPSPAPNRHSRQAPNVAQELSGGCHGSHKPMHWQEBLLGG 301
QY 566 -----EAEALIMTF-----SLNNYPAGDPRLAQAOKLWERAFLERMAFORRMAG 609
D 302 TARDLQQLLRAEALQSTELMSPROLYEHFRGDYQTHDIGNSEQAQSMVLQAWQRRFVQ 361
QY 610 MEQVFTTAERSLE-----DEINRTTAEDLPFATSYIV--IFLYISLALGSYSWS 658
D 362 LAQEAALPANASQOIHFASFTTLLDILRAFSE-----VSTTRVVGGLMLAYACVTMLRWD 417
QY 659 RVMDSKATLGLGGVAVLGVMAAMGPFPSYILGIRSSVLQVVPFLVLSVGADNIFIV 718
D 418 --CAOSQAGVLAGVLLVALAVASGLGALLGTFNAATTQVLPFALGIGVDDIFLLA 475
QY 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEACFFFLGALTMPAVRTFALTSGL 778
D 476 HAFTRAP--PDTPERPMEGCLRTGTSVALTSVNNMVAFFMAALVPFIPALRAPSLQAAI 533
QY 779 AVILOFLLQMSAFVALLSDSRQEARSLDVCC-----VKQOE----- 817
D 534 VVGCNFAAVMLVFPAILSLDLRRRHRQRLDLVLCFSSPCSAQVIOQLPOELGDRAPVGI 593
QY 818 -----LBP-----PG-----QEGE-- 826
D 594 AHLTATVQAFTHCBASSQHVWTLPPQAHLLSPASDPLGSELYSPGGSTRDILLSBEGTG 653
QY 827 -----LLGPFQKAYAPLHLWITRGVLLFLFALFGVSLYSMCHISVGLDQ 873
D 654 POACRPLLCALHTLAHPARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713
QY 874 ELALPKDSVLLDYFLNRYFEVGAPVYFVTLGNFNS--SEAGMAICSSACNNFSTQ 932
D 714 TDVVPRTKEHAFSLAQLRYSFL--YEVALVTQGFDAHSQAL-----FDLHQ 761
QY 933 KIQYATEPPEQSYLAIPAS-----SWDDF-----JDLTPSSCCRLYISGPNK 976
D 762 RFSSL-----KAVLPPATQAPRTWLHYTRSLQGIQAADFQDQWASGRITCHSYRNGSD 816
QY 977 D-----KFCPSTVNSLNLKNCMSITM-----GSVRPSVEQFHKYLPFPLNDRPNKCPK 1026
D 817 GALAYKLLIQTGNAQEPLDPSQLTTRKLVKESGLIPP--ELFYMGLTVWVSSDPL----- 869
QY 1027 GGIAA-----YST-----SVNLTSQGVLASRPMAYHKPLKNSQDYTEALR 1067
D 870 -GLAASQANFYPPPEWLHDKDVTGTENLRIPAAQPLEPAQFPFLLHGLQKTADPVEAIE 928
QY 1068 ABELAANI--TADLRKVPCTDPAFVFPYTIINVFYEQVLTLPBGLFMLSIC--LVPTFF 1124
D 929 GARAACTEAGQGVHAYPSGSPF-----LFWEOYLGV--RRCFLAVCILLVCTVF 976
QY 1125 AVSCLLGLDLRSLNLILSIWMILVDVTGFMALWDISYNVSLNLINVASVGMSEVFFVSH 1184
D 977 LVCALLLSPWTAGLI--VLVLAMVTVELFGMGFIGIKLSAIPVILVASIGIVEFTWH 1035
QY 1185 ITRSPAISTKPTWLBRKAEATISMGSAVPAGV---AMTNLPGLVILGLAKAQLIOIPFR 1241
D 1036 VALGELTSHGSRNLRAA-----SALEQTPAPVTDGAVSTLLGLMLAGSNFDFIIRYFV 1090
QY 1242 LNLTLTLGLLGLHGLVPLVILSVGP 1267
D 1091 VLTVTLTLGLLGLHGLLPLVILSILGP 1116
```

RESULT 14

```
US-11-337-244-149
; Sequence 149, Application US/11337244
```

GENERAL INFORMATION:

```
; APPLICANT: William Matthews
; APPLICANT: Mark Moore
; APPLICANT: Russell Phillips
; APPLICANT: Michael V. Wiles
; APPLICANT: Thadd C. Reeder
; APPLICANT: Robert G. Wisotzkey
; APPLICANT: Keith D. Allen
; APPLICANT: Helen Baribault
; APPLICANT: Thomas J. Brennan
; APPLICANT: Catherine Guenther
; APPLICANT: Robert Klein
; APPLICANT: Christopher J. Kirk
; APPLICANT: Simon X. Xie
; APPLICANT: Qin Zhang
; APPLICANT: Agnes Choppin
```

```
; TITLE OF INVENTION: Targeted Gene Disruptions, Compositions
; FILE REFERENCE: R-DB-13
```

```
; CURRENT APPLICATION NUMBER: US/11/337,244
; CURRENT FILING DATE: 2006-01-20
```

```
; PRIOR APPLICATION NUMBER: US/10/669,143
```

```
; PRIOR FILING DATE: 2003-09-23
```

```
; PRIOR APPLICATION NUMBER: US 60/413,647
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: US 60/413,666
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: US 60/413,653
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: US 60/413,646
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: US 60/413,625
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; PRIOR APPLICATION NUMBER: US 60/413,639
```

```
; PRIOR FILING DATE: 2002-09-24
```

```
; NUMBER OF SEQ ID NOS: 237
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 149
```

```
; LENGTH: 1182
```

```
; TYPE: PRT
```

```
; ORGANISM: Mus musculus
```

```
US-11-337-244-149
```

Query Match

```
8.4%; Score 583.5; DB 6; Length 1182;
```

```
Best Local Similarity 24.2%; Pred. No. 4.7e-38;
```

```
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;
```

```
QY 330 SPSHTLIGQ-----PFQG-----WGTWVASPLTILVLSVIPVVALAAGLVFTLT 376
D 20 SSAPHITAGSQAPLWLRAYFQGLFSLGCRQKHGKVLFLGLVAFGALGLRAVIE 79
QY 377 TDPVELWAPNSQARSEKAFDQHPG--PFFRTNQVILTAPNRSSVRYDSLLGPKNFSGI 435
D 80 TDLEQLWVEGSRVSOELHYHKEKLGEEAANTYSQLI-----QTAHQEGNVLTPE---A 131
QY 436 LDLDLLELLELQERLHLOVNSPEAQRNISIQIDICY---APLNPDN-----TSLYDCC 486
D 132 LD-----LHLQAALTASKQVSVYKSGWDLNKKCYKSGVPLIENGMIERMIKLFPCV 184
QY 487 INSLLOYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---PKDGTALA 539
D 185 ILTPLDCFWEGAK---LOGGSAYLPGRPDIOQNTLNDPQOLLEELGPFASLEGFRELDDKA 241
QY 540 LSCMADYGAPVF-----PFLA---IGYKKGKDYs----- 565
D 242 QVGQAYVGRPCLDPPDPPHCPSPAPNRHSRQAPNVAQELSGGCHGSHKPMHWQEBLLGG 301
QY 566 -----EAEALIMTF-----SLNNYPAGDPRLAQAOKLWERAFLERMAFORRMAG 609
```

Db 302 TARDLOGQLRAEALQSTFLMSPRQLYEHFRGDYQTHDIGWSEBQASMWLQAWQRRFVQ 361
Qy 610 MPQVTTFAERSLE-----DEINRTAEDLPFATSYIV--IFLYISIALGYSYWS 658
Db 362 LAQEALPANASQOIHAFFSTTDDILRAFSE-----VSTRVVGGLMLAYACVTMLRWD 417
Qy 659 RVWVDSKATLGGVAVVAGVMAAGFFSYLGRSSVLIVQVPPFLVSLVSGADNIFV 718
Db 418 --CAQSGAVGLAGVLLVALAVASGLICALLGITFNATQVLPFLGALGVDIFLLA 475
Qy 719 LEYQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEACFFLGALTMPAVRTPTALTSGL 778
Db 476 HAFTKAP--PDTFLPERMGECLRTGTSTGTSVLTNNVMAFFMAALVPALRAFSLQAAI 533
Qy 779 AVILDFLOMSAFVALLSLSKQASRLDVCCC-----VKQOE----- 817
Db 534 VVGCNFAAVMLVFPAILSLDLRRRRQRDLVLCFSPSCSAQVIQMLPQELGDRAPVGI 593
Qy 818 -----LPP-----PG-----QEGG-- 826
Db 594 AHLTATVQAFTHCEASSQHVWITLPQAHLLSPASDPGLGSELYSGGSTRLDLSQEGTG 653
Qy 827 -----LLLGFPQKAYAPFLHMITRGVWVLLFLFALFGVLSYMSCHISVGLDQ 873
Db 654 PQAACRPLLCAHWTLAHFARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713
Qy 874 ELALPKOSVLLDYFLNRYFVGAPVYFVTTLGYNFS--SEAGMAICSSACGNNFSTQ 932
Db 714 TDVPRGTKEHAFSLAQLEYFSL--YEVALVTOGDFDYAHSORAL-----FDLHQ 761
Qy 933 KIOYATEFPEQSVLAIPAS-----SWDDF-----IDMLTPSSCCRLYISGPNK 976
Db 762 RFSSL-----KAVLPPTAQAPRTWLHYRSLWLOGIQAAFDQDQWASGRITCHSYNGSED 816
Qy 977 D-----KFCPSTVNSLNCNCHSITM-----GSRVPSVEQPHKYLPMFLNDRPNKCPK 1026
Db 817 GALAYKLLIQTGNAQBPDFSQLTTRKLVDEKGLIPP--ELFVGMGLTVWVSDPL----- 869
Qy 1027 GGLAA-----YST-----SVNLTSQGVLASRFWAYHKLKNSQDYTEALR 1067
Db 870 --GLAASQANFYPPPEMLHDKYDTTGENLRIPAAQPLEFAQFPFLHGLQKTAQDFEATE 928
Qy 1068 AARELAANI--TADLRKVPGTDAFAVEFPYITNVFYEQYLITLPEGLFMLSIC--LVPTF 1124
Db 929 GARAACTEAGQAGVHAYPSGSPF-----LFWEQVGL--RECFLAVCILLVCTF 976
Qy 1125 AVSCLLGLDLRGLLNLISVIMLVDTVTGFMALWDISYNAVSLINLVSAGVMSVEFVSH 1184
Db 977 LVCALLLSPTAGLI--VLVAMMTVELFGIMFGILGKLSAIPVVLVASIGIGVBEFTVH 1035
Qy 1185 ITRSFALSTKPTWLERAKEATISMGSAVFAGV--AWTNLPGLVLGLAKAQLIQIFFR 1241
Db 1036 VALGFUTSHGSRNLRAA-----SALSGTAPVTDGAVSTLLGLLMLAGSNFDFIRYFV 1090
Qy 1242 LNLTLGLLGLHGLVFLPVILSVXP 1267
Db 1091 VLTVTLGLLGLHGLLPLVLLSILGP 1116

RESULT 15

US-60-772-265-1197

; Sequence 1197, Application US/60772265

; GENERAL INFORMATION:

; APPLICANT: BOUKHAROV, ANDREY A.

; APPLICANT: DU, ZIJIN

; APPLICANT: GUO, LIJANG

; APPLICANT: HRESKO, MICHELLE C

; APPLICANT: KOVALIC, DAVID K

; APPLICANT: ZHAOLONG, LI

; APPLICANT: LU, MAOLONG

; APPLICANT: MCCARTER, JAMES P

; APPLICANT: MILLER, NANCY M

; APPLICANT: VAUDIN, NANCY

; APPLICANT: WILLIAMS, DERYCK J
; APPLICANT: WU, WEI
; TITLE OF INVENTION: IDENTIFICATION AND USE OF TARGET GENES
; FILE REFERENCE: MNDI:002USP1
; CURRENT APPLICATION NUMBER: US/60772,265
; CURRENT FILING DATE: 2006-02-10
; NUMBER OF SEQ ID NOS: 1919
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1197
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: genomic DNA sequence=SeqID_396; coding sequence=SeqID_981
US-60-772-265-1197

Query Match 7.0%; Score 482; DB 8; Length 891;

Best Local Similarity 22.2%; Pred. No. 5,3e-30;

Matches 203; Conservative 160; Mismatches 36; Indels 190; Gaps 35;

Qy 540 LSCWADYCAPVPFPLAIGGYK---GKDYSEAEALIMTFSLNNYPAGD--PRLAQAK--L 591
Db 10 MNCPEP-----LIVGGIRRDAGQVIRHAEAFQTVFLVAS--AGDVQRLRSKNHR 59
Qy 592 WEEAFLEBEMRAFQRRMAGMFQVTTAERSLEDEINRTTAEDL----PIFATS----- 639
Db 60 MEKGVSTQAHAGAITAAWQNF--KSIYDHTLNKPAQGIRVVVHPLASTSIQDMLRQF 117
Qy 640 ----YVIFL--YISLALGYSYWSRV-----WVDSKATLGLGVAVVAGVMAAGFF 687
Db 118 SEQPFVIFIGVLMII--YAGWSQVHWQGMFYSKSCLLAIIGVLVITLASVAGLGUS 175
Qy 688 SYLGIRSLVILQVVPFLVLSVGADNIFVLEYQRLPRRCPPEPREVHIGRALGRVAPSM 747
Db 176 TAMNHFNAATTQIVPFLTGLGDDMFLLHNYDVLE--AVRQKEVAVLLKETGMSV 232
Qy 748 LCLSEIAICFFLGALTMPAVRTFALTSLGLAVILDPLQLMSAFVALLSLSKQAEARL 807
Db 233 LITSINNILAFITGCILPILALRSFGQVAILTNSVLCIILLPFAITALLRRKAGHR 292
Qy 808 DVCCC-----VKQELPPQGE-----GALLGFFQKAVAP 838
Db 293 DMSFCSSRNQLVKTDSTVTHSSDELDQMSAAVETVSATKSLHKWTLVGLHGYIP 352
Qy 839 FLHWHITGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGA 898
Db 353 LLRRPIAKTVVLLVCAAMFLFCGFLYYSRIGLELADVLPEHTAPAAFLKARERYFSP-Y 411
Qy 899 PVYFV---TTLGYNFSSEAGMAICSSACGNNF-----SFTQKI 934
Db 412 PMFIVPKGPMVDY--ANQHQKLEQLRARDIGRSNFVIKVDGEPQWMMFMRTWLHLSQSSL 470
Qy 935 QVATE--FPQSYLAI PASSWV-----DDFDWLTPTSSCCRLYISGNKDKFCPTVN 985
Db 471 DRAAKIGIPDDDLKXMAADALARNFTLSDDPL--LARKLLC-----SAQ 514
Qy 986 SLNC--LKNCMSITMGSVRPSVEQPHKYL--PMFLNDRPNKCPKGG-----LAAYS 1033
Db 515 RENCTEMGAKLIEDGRINP--RGFNYLTAFYQDNMMYVYSQAFFPTPLPWHPFSAAD 572
Qy 1034 TSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDAFAVEF 1093
Db 573 ESVVPPAD--PLLYSQIPFYMNGLTDTQSIQVQMIKEIRAICDRYSAD-----GLPVY 622
Qy 1094 PTTITNVFEQYLTILPEGLFMLSICLVPTFAVSCLLGLDLRLSGLLNLISLTVMI--LYD 1151
Db 623 PSGIPPTFWEQYKLT---FYLVSAILIIGVAVLLVISIIIFNPWAAAMVAIVTWVTE 679
Qy 1152 TVGFNALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSA 1211
Db 680 LAGFMGVGVGNPISAVTLITAVGIVEFTAVHVLAFITS-----LGRSDRMWACLEH 734

THIS PAGE IS BLANK

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: April 11, 2006, 00:36:06 ; Search time 201 Seconds
 (without alignments)
 2911.704 Million cell updates/sec

Title: US-10-736-769-4
 Perfect score: 6909
 Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues
 Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 100 summaries

Database : A_Geneseq_21.*
 1: Geneseqp1980s.*
 2: Geneseqp1990s.*
 3: Geneseqp2000s.*
 4: Geneseqp2001s.*
 5: Geneseqp2002s.*
 6: Geneseqp2003as.*
 7: Geneseqp2003bs.*
 8: Geneseqp2004s.*
 9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	8	ADJ27297 Human NPC
2	6909	100.0	1332	9	ADY60740 Human Nle
3	6909	100.0	1332	9	AEB93570 Human NPC
4	6896	99.8	1332	4	AAM79169 Human pro
5	6896	99.8	1332	4	AAG65638 SSD-conta
6	6872.5	99.5	1359	8	AAM79168 Human pro
7	6872.5	99.5	1359	8	ADJ27337 Human NPC
8	6872.5	99.5	1359	9	ADY60780 Human Nle
9	6872.5	99.5	1359	9	AEB93610 Human NPC
10	6536	94.6	1344	4	AEB22593 Novel hum
11	5421.5	78.5	1331	8	ADJ27295 Rat NPCIL
12	5421.5	78.5	1331	9	ADY60738 Rat Niema
13	5421.5	78.5	1331	9	AEB93568 Rat NPCIL
14	5407	78.3	1333	8	ADJ27305 Mouse NPC
15	5407	78.3	1333	9	ADY60748 Mouse Nle
16	5407	78.3	1333	9	AEB93578 Mouse NPC
17	4466	64.6	982	4	ABG22691 Novel hum
18	2402.5	34.8	1278	3	AAB88445 Human NPC
19	2402.5	34.8	1278	3	AAB42983 Human ORF
20	2402.5	34.8	1278	8	ADQ39879 Human myo
21	2402.5	34.8	1278	8	ADU06723 Novel bro
22	2385	34.5	1319	2	AAW88446 Mouse NPC
23	2262	32.7	1287	4	AB61737 Drosophil
24	2262	32.7	1287	8	ADS96670 Drosophil

25	1864.5	27.0	1223	4	ABB58629	Abb58629 Drosophil
26	1329	19.2	1170	2	AAW88447	Aaw88447 Yeast NPC
27	1078	15.6	1296	2	AAW88448	Aaw88448 Caenorhab
28	963	13.9	229	4	AAM80153	Aam80153 Human pro
29	963	13.9	229	4	AAM80152	Aam80152 Human pro
30	942	13.6	194	4	AAG65637	Aag65637 SSD domai
31	651.5	9.4	1447	2	AAR75375	Aar75375 Human pat
32	651.5	9.4	1447	2	AAW52200	Aaw52200 Human pat
33	651.5	9.4	1447	2	AAW72969	Aaw72969 Human pat
34	651.5	9.4	1447	2	AAB67163	Aab67163 Human pat
35	651.5	9.4	1447	5	AAE19830	Aae19830 Human pat
36	651.5	9.4	1447	5	ABU10931	Abu10931 TRC8 rela
37	651.5	9.4	1447	5	AAG79571	Aag79571 Human pat
38	651.5	9.4	1447	7	ABU62275	Abu62275 Human pat
39	651.5	9.4	1447	7	ADD46678	Add46678 Human Pro
40	651.5	9.4	1447	7	AD894224	Ad894224 Human ptc
41	651.5	9.4	1447	7	ADH62731	Adh62731 Human pat
42	651.5	9.4	1447	8	AD848989	Ad848989 Human pat
43	646.5	9.4	1434	2	AAW52199	Aaw52199 Mouse pat
44	646.5	9.4	1434	2	AAW72968	Aaw72968 Mouse pat
45	646.5	9.4	1434	4	AAB67159	Aab67159 Murine pa
46	646.5	9.4	1434	5	AAG79572	Aag79572 Mouse pat
47	646.5	9.4	1434	7	ABU62271	Abu62271 Mouse pat
48	646.5	9.4	1434	7	ADH62722	Adh62722 Mouse pat
49	646.5	9.4	1434	8	AD848980	Ad848980 Mouse pat
50	646.5	9.4	1434	8	AAR94380	Aar94380 Mouse pat
51	643.5	9.3	1434	2	AAW47157	Aaw47157 Nevoid ba
52	628	9.1	1296	7	ADJ21144	Adj21144 Novel hum
53	588.5	8.5	954	5	ABB98519	Abb98519 Human tes
54	588	8.5	954	5	AAE19829	Aae19829 Human pat
55	584	8.5	1203	5	AAE19829	Aae19829 Human pat
56	584	8.5	1203	5	ABG74104	Abg74104 Human pat
57	584	8.5	1203	6	ABG74104	Abg74104 Human pat
58	583.5	8.4	1182	5	AAE19831	Aae19831 Mouse pat
59	583	8.4	1203	3	AAE19831	Aae19831 Mouse pat
60	580	8.4	1203	3	AAE19831	Aae19831 Mouse pat
61	579	8.4	1146	3	AAE19831	Aae19831 Mouse pat
62	531	7.7	1061	4	ABB65553	Abb65553 Drosophil
63	526.5	7.6	1311	2	AAW72971	Aaw72971 Precis co
64	526.5	7.6	1311	4	AAE67156	Aae67156 Butterfly
65	526.5	7.6	1311	5	AAE67156	Aae67156 Butterfly
66	526.5	7.6	1311	5	ABU62149	Abu62149 Buckeye p
67	526.5	7.6	1311	7	ADH62716	Adh62716 Peacock b
68	526.5	7.6	1311	7	ADH62716	Adh62716 Peacock b
69	526.5	7.6	1311	8	AD848974	Ad848974 Butterfly
70	514	7.4	1405	8	ADN22811	Adn22811 Bacterial
71	512.5	7.4	1311	2	AAW52197	Aaw52197 Precis co
72	474	6.9	1286	4	ABB59092	Abb59092 Drosophil
73	474	6.9	1286	5	ABJ10929	Abj10929 TRC8 rela
74	464.5	6.7	1299	4	AAE05389	Aae05389 Human pat
75	464.5	6.7	1299	4	AAE05389	Aae05389 Human pat
76	464.5	6.7	1299	4	AAE05389	Aae05389 Human pat
77	464.5	6.7	1299	4	AAE05389	Aae05389 Human pat
78	464.5	6.7	1299	5	ABJ1232	Abj1232 A consens
79	464.5	6.7	1299	5	ABJ1232	Abj1232 A consens
80	464.5	6.7	1299	6	ADA26289	Ada26289 Hedghog-
81	464.5	6.7	1299	7	ADN25309	Adn25309 Hedghog-
82	464.5	6.7	1299	7	ABW00878	Abw00878 Drosophil
83	464.5	6.7	1299	7	ABW01685	Abw01685 Drosophil
84	464.5	6.7	1299	7	ADH61129	Adh61129 Drosophil
85	464.5	6.7	1299	8	ADH54692	Adh54692 Drosophil
86	464.5	6.7	1299	8	ADU26454	Adu26454 Drosophil
87	464.5	6.7	1299	9	ADW25903	Adw25903 Drosophil
88	464.5	6.7	1299	9	ADW25903	Adw25903 Drosophil
89	438.5	6.3	1285	2	AAW72972	Aaw72972 Fruit fly
90	438.5	6.3	1285	4	AAW72972	Aaw72972 Fruit fly
91	438.5	6.3	1285	5	AAW72972	Aaw72972 Fruit fly
92	438.5	6.3	1285	7	ABU62150	Abu62150 Fruit fly
93	438.5	6.3	1285	7	AD894211	Ad894211 Fruit fly
94	438.5	6.3	1285	7	ADH62718	Adh62718 Fly patch
95	379	5.5	714	8	ADN22757	Adn22757 Drosophil
96	369.5	5.3	803	8	ADQ66737	Adq66737 Bacterial
97	351.5	5.1	767	5	ABB98520	Abb98520 Human tes

98 349 5.1 783 7 ADI21155 Ad21155 Novel hum
99 339 4.9 783 6 ADA54204 Ada54204 Human pro
100 316.5 4.6 608 7 ADD46676 Add46676 Rat Prote

ALIGNMENTS

RESULT 1
ADJ27297
ID ADJ27297 standard; protein; 1332 AA.

AC ADJ27297;
DT 20-MAY-2004 (first entry)
XX Human NPC1L1.
XX Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

OS Homo sapiens.
XX WO2004009772-A2.
XX 29-JAN-2004.
XX 17-JUL-2003; 2003WO-US022467.
XX 19-JUL-2002; 2002US-0397442P.
XX (SCHE) SCHERING CORP.
XX Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2004-132945/13.
XX N-PSDB; ADJ27296.
XX GENBANK; AF192522.

XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
PT or identifying antagonists of NPC1L1 for inhibiting intestinal
PT cholesterol absorption in a subject, or for treating elevated serum
PT cholesterol or stroke.

XX Claim 11; SEQ ID NO 4; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1
CC (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
CC a motif ADJ27331 which acts as a trans-golgi network to plasma membrane
CC transport signal, and which exhibits limited tissue distribution and
CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a
CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels,
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42% amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.

XX Sequence 1332 AA;

Query Match 100.0%; Score 6909; DB 8; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1081 RKVPGTDPAPFPPYPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAPFPPYPTITNVFYEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIWMLIVDTVGFMAWDISTNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMLIVDTVGFMAWDISTNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVAGVAMTNLPGILVGLAKAQLIQIFFFRNLMLITTLGLHLGLVFLPV 1260
Db 1201 AKEATISMSGSAVAGVAMTNLPGILVGLAKAQLIQIFFFRNLMLITTLGLHLGLVFLPV 1260
QY 1261 ILSYVGPDPNPALALEQKRAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPNPALALEQKRAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332

RESULT 2
ADY60740
ID ADY60740 standard; protein; 1332 AA.
XX ADY60740;
AC XX
19-MAY-2005 (first entry)
DE Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 4.
XX
KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX
OS Homo sapiens.
XX WO2005015988-A1.
XX
PD 24-FEB-2005.
XX
PF 16-DEC-2003; 2003WO-US040113.
XX
PR 17-JUL-2003; 2003US-00621758.
PR 22-AUG-2003; 2003US-00646301.
PR 16-SEP-2003; 2003US-00663208.
XX
PA (SCHE) SCHERING CORP.
XX
PI Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
DR N-PSDB; ADY60739, ADY60787.
XX
XX Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
XX
PS Claim 30; SEQ ID NO 4; 146pp; English.
XX
CC The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the

CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist,
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of
CC ezetimibe or substituted azetidinone), decreasing the level of intestinal
CC sterol or Salpha-sterol absorption in a subject (involving reducing the
CC level of expression of NPC1L1 in the subject), identifying an antagonist
CC of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
CC functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
CC of NPC1L1 which is useful for inhibiting or decreasing the level of
CC NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
CC mouse, rat or human. The level of expression of NPC1L1 in the subject is
CC reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
CC useful for screening a sample for intestinal sterol or Salpha-sterol
CC absorption antagonist. The NPC1L1 antagonists identified are useful for
CC treating medical conditions caused or mediated by NPC1L1, e.g.,
CC hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or
CC arteriosclerosis. The present sequence represents a mammalian NPC1L1
CC protein.
XX
SQ Sequence 1332 AA;
Query Match 100.0%; Score 6909; DB 9; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLGRWLLWALLRLAQAQSEPTTTHQPGVCAFYDECGKPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLGRWLLWALLRLAQAQSEPTTTHQPGVCAFYDECGKPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLQKICPRLYTGNTQACSAKQVLSLEASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLQKICPRLYTGNTQACSAKQVLSLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSDSCSRVPA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVAYEAPYQHSFAEQSDSCSRVPA 180
QY 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAFLDITFHLLEPGQAVSGIQLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAFLDITFHLLEPGQAVSGIQLNEGV 240
QY 241 ARCNEQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVETELTDPVELWSAPNSQARSEKAFPHDQHPFRFTNQVILTPNRSY 420
Db 361 IPVVALAAGLVETELTDPVELWSAPNSQARSEKAFPHDQHPFRFTNQVILTPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLDDLLLELQERLHLQVWSPQAORNTSLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLLLELQERLHLQVWSPQAORNTSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLLLTANQTLMGQTSQVDMKQHFYCANAPLTFKDGFTAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLLLTANQTLMGQTSQVDMKQHFYCANAPLTFKDGFTAL 540
QY 541 SCMADYGAPVPPFLAIGYKGYKDYSEALINTFSLNNYPAGDPRLAOKLWEAFLEEM 600
Db 541 SCMADYGAPVPPFLAIGYKGYKDYSEALINTFSLNNYPAGDPRLAOKLWEAFLEEM 600
QY 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSVR 660
Db 601 RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPFPATSYIVIFLYISALGYSWSVR 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSSLVLQVVPFLVLSVGADNIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSSLVLQVVPFLVLSVGADNIFVLE 720

Db 661 MVDKATLGLGVAVVLGAVNAAGFFSYLGIRSLVLQVVPFLVLSVGDNIFFIVLE 720
QY 721 YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTMPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTMPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQWSAFVALLSLDSKQBSRLDVCCKVPOELPPQGGEGLLIGFFQKAYAPFL 840
Db 781 ILDFLLQWSAFVALLSLDSKQBSRLDVCCKVPOELPPQGGEGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
Db 841 LHWITRGVVLLLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
QY 901 YFVTTLYGNSFSEAGMNAICSSAGCNFFSFTQKIYATEFFPEQSYLAIPASSWDDDFIDW 960
Db 901 YFVTTLYGNSFSEAGMNAICSSAGCNFFSFTQKIYATEFFPEQSYLAIPASSWDDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITWGSVRPSVEQFHXYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITWGSVRPSVEQFHXYLPWFLNDRP 1020
QY 1021 NIKCPKGGIAAYSTVNLTSQVLSRPMAYHKLPKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGGIAAYSTVNLTSQVLSRPMAYHKLPKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
Db 1081 RKVPGTDPAPFVPPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGILL 1140
QY 1141 NLLSIWMLVDTVGFMAWDISVNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFMAWDISVNAVSLINLVSAGVMSVEFVSHITRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGDVNPALALQKRAEBAVAAMVASCNHPHSRSTADNIYVNHSPSGSIKGAGA 1320
Db 1261 ILSYVGDVNPALALQKRAEBAVAAMVASCNHPHSRSTADNIYVNHSPSGSIKGAGA 1320
QY 1321 ISNFPNNGRQF 1332
Db 1321 ISNFPNNGRQF 1332
RESULT 3
AEB93570
ID AEB93570 standard; protein; 1332 AA.
AC AEB93570;
XX
DT 06-OCT-2005 (first entry)
XX
DE Human NPC1L1 protein.
XX
KW Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;
KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
KW arteriosclerosis; human.
OS Homo sapiens.
XX
PN W02005069900-A2.
XX
PD 04-AUG-2005.
XX
PF 14-JAN-2005; 2005WO-US001469.
XX
PR 16-JAN-2004; 2004US-0537341P.
XX (MERI) MERCK & CO INC.
XX

PI Garcia-Calvo M;
XX WPI; 2005-564070/57.
DR N-PSDB; AEB93569.
XX
PT Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by
contacting NPC1L1 with detectably labeled substituted 2-azetidinone
glucuronide and a candidate compound and determining if compound binds to
human NPC1L1.
XX
PS Example 2; SEQ ID NO 4; 215pp; English.
XX
CC The invention relates to identifying a ligand of NPC1L1. The method
involves contacting human NPC1L1 with a detectably labeled substituted 2-
azetidinone glucuronide and a candidate compound and determining whether
the candidate compound binds to human NPC1L1. In identifying a ligand of
NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands
are useful for stimulating or blocking the activity of NPC1L1, and for
treating conditions caused or mediated by NPC1L1. It is useful for
reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
disease, stroke, or arteriosclerosis. The present sequence represents a
human NPC1L1, a N-glycosylated protein.
XX
SQ Sequence 1332 AA;
Query Match 100.0%; Score 6909; DB 9; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECCKNPGLSGMLTSLNVCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTIHQPGYCAFYDECCKNPGLSGMLTSLNVCLSN 60
QY 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKOLVSLASLSITKALLTRCPACSNF 120
Db 61 TPARKITGDHLLILQKICPRLYTGNTQACCSAKOLVSLASLSITKALLTRCPACSNF 120
QY 121 VNLHCHNTCSNQSILFINVTRVAOLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSILFINVTRVAOLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCGVYGSAALCNAQRWLNFGQDGTGNGLAPLDTIFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSAALCNAQRWLNFGQDGTGNGLAPLDTIFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQGGDVATCSCQDCNAACSCPAIARPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNESQGGDVATCSCQDCNAACSCPAIARPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGFRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQHFPPFTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQHFPPFTNQVILTAPNRSSY 420
QY 421 RYDSLILGPKNFSGILDLILLLELLELQERLRLHQLQWSPEAQKNISLQDICYAPLNPNT 480
Db 421 RYDSLILGPKNFSGILDLILLLELLELQERLRLHQLQWSPEAQKNISLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANTQTLMGQTSQVDWKDHFLYCANABLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANTQTLMGQTSQVDWKDHFLYCANABLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKDYSEAEALIMTFESLNNYPAGDPRIQAOKLWEAEFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKDYSEAEALIMTFESLNNYPAGDPRIQAOKLWEAEFLEEM 600
QY 601 RAFOREMAGMFOVTFTAERSLEDEINRTTAEDLIPFATSYIVIFLYISALGYSWSRV 660
Db 601 RAFOREMAGMFOVTFTAERSLEDEINRTTAEDLIPFATSYIVIFLYISALGYSWSRV 660

Db 601 RAQREMGAFQVTTAERSLEDEINRTAEDLPFATSIVIVFLYISIALGSYSWSRV 560
Qy 661 MYDSKATLGGVAVVVGAAVMAAMGPFYSGIRSSLVILQVVPFLVLSVGCADNIFVILE 720
Db 661 MYDSKATLGGVAVVVGAAVMAAMGPFYSGIRSSLVILQVVPFLVLSVGCADNIFVILE 720
Qy 721 YORLPRPGEPRVHIGRALGRVAPSMILCSISEAICFFLGALTMPAVRTTALTSLAV 780
Db 721 YORLPRPGEPRVHIGRALGRVAPSMILCSISEAICFFLGALTMPAVRTTALTSLAV 780
Qy 781 ILDFLLQMSAFVALLSDSKREASRLDVCCCKVQELPPPGQEGELGFGKAVAPFL 840
Db 781 ILDFLLQMSAFVALLSDSKREASRLDVCCCKVQELPPPGQEGELGFGKAVAPFL 840
Qy 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPV 900
Qy 901 YFVTTLYGNFSSEAGNNAICSSAGCNFSTQKIYATEPPEQSYLAIPASSWVDDFDIW 960
Db 901 YFVTTLYGNFSSEAGNNAICSSAGCNFSTQKIYATEPPEQSYLAIPASSWVDDFDIW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCKMSITMGSVRPSVSEPHKYLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCCKMSITMGSVRPSVSEPHKYLWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFVFPYITNVFTEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFVFPYITNVFTEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLL 1140
Qy 1141 NLLSIVMILVDVTGFMALWDISYNAVSLNLVSAVGMSEVFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDVTGFMALWDISYNAVSLNLVSAVGMSEVFVSHITRSPAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLTITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLTITLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVNPALALEOKRAEBAVAAMVASCNPHSPRVSTADNIYVNHSPSGSIKGAGA 1320
Db 1261 ILSYVGPVNPALALEOKRAEBAVAAMVASCNPHSPRVSTADNIYVNHSPSGSIKGAGA 1320
Qy 1321 ISNLPNNGRQF 1332
Db 1321 ISNLPNNGRQF 1332

RESULT 4

AAW79169

ID AAW79169 standard; protein; 1332 AA.

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX (HYSSE-) HYSEQ INC.
XX Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;
XX WPI: 2001-476283/51.
DR N-PSDB; AAK52302.
XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
in diagnosis and gene therapy.
XX Claim 20; Page 4216-4219; 5221pp; English.
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
cytokine, cell proliferation or cell differentiation or which may induce
production of other cytokines in other cell populations. The
polynucleotides and polypeptides are useful in gene therapy, vaccines or
peptide therapy. The polypeptides have various cytokine-like activities,
e.g. stem cell growth factor activity, haematopoiesis regulating
activity, tissue growth factor activity, immunomodulatory activity and
activin/inhibin activity and may be useful in the diagnosis and/or
treatment of cancer, leukaemia, nervous system disorders, arthritis and
inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
(AAK52582) and 3666 (AAK80020) are omitted as the relevant pages from the
sequence listing were missing at the time of publication
XX SQ Sequence 1332 AA;
Query Match 99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAPYDECKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAPYDECKNPELSGLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSIFINVTVAQOLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
Db 121 VNLHCHTCSNQSIFINVTVAQOLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRVPA 180
Qy 181 ATLAVGTGCVYGSALCNAQRWLNFGDGTNGCLAPLDITFHLLEPQAVGSGIQLNEGV 240
Db 181 ATLAVGTGCVYGSALCNAQRWLNFGDGTNGCLAPLDITFHLLEPQAVGSGIQLNEGV 240
Qy 241 ARCNSQGDVATCSQDCAASCPAIPALQALDSTFYLGQMPGSLVLIILCSFAVVTI 300
Db 241 ARCNSQGDVATCSQDCAASCPAIPALQALDSTFYLGQMPGSLVLIILCSFAVVTI 300
Qy 301 LLVGRVAPARDKSKQVDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKQVDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTDTPVELWSAPNSQAKSEKAFHQHFGFPFRFTQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTLTDTPVELWSAPNSQAKSEKAFHQHFGFPFRFTQVILTAPNRSY 420
Qy 421 RYDSLLLGPKNFGSLDLDLLELLELQERLRLHQLVWSPEAQRNISLQDICYAPLNDPT 480

Db 421 RYDSELLGPKNFSGTLDLLELLELQERLRLHQLVMSPEAQRLNISLQDICYAPLNDPT 480
Qy 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGTSQVDWKDHLFYCANAPLTFKDGTLAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGTSQVDWKDHLFYCANAPLTFKDGTLAL 540
Qy 541 SCMDYGAVPFPFLAIGYKGVKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
Db 541 SCMDYGAVPFPFLAIGYKGVKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEM 600
Qy 601 RAFORMAGMFOVTTAERSLEDEINRTTADLPFATSYIVFLYISLALGSSYSSSRV 660
Db 601 RAFORMAGMFOVTTAERSLEDEINRTTADLPFATSYIVFLYISLALGSSYSSSRV 660
Qy 661 MYDSKATLGLGCVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIPFVLE 720
Db 661 MYDSKATLGLGCVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIPFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGNALTPMPAVRTFALTSLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLGNALTPMPAVRTFALTSLAV 780
Qy 781 ILDELLQMSAFVALLSLDSKROEASRLDVCCCKQELPPQCGGLLGFQKAYAPFL 840
Db 781 ILDELLQMSAFVALLSLDSKROEASRLDVCCCKQELPPQCGGLLGFQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFVGPAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFVGPAPV 900
Qy 901 YFVTTLYGNPFSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDPDFDW 960
Db 901 YFVTTLYGNPFSAGNNAICSSAGCNFSTQKIQYATEFPQSYLAIPASSWVDPDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNCSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNLCKNCSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTS DGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTS DGQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAPFVFPYITNVFYEQYLTILPEGLFMLSCLVPTFVNSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAPFVFPYITNVFYEQYLTILPEGLFMLSCLVPTFVNSCLLGLDLRSGLL 1140
Qy 1141 NLLSIWMILVDTVGFMAWLDISYNAVSLNLVSAVGMVSEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIWMILVDTVGFMAWLDISYNAVSLNLVSAVGMVSEFVSHITRSPAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVPAVAMTNLPGILVLGLAKAQLIQIPFRLNLLITLGLLHGLVFLPV 1260
Qy 1261 ILSVGDVNPALAEQKRAEEAAVAVMVASCPNHPRSVSTADNIYVNSHFEKSIKAGA 1320
Db 1261 ILSVGDVNPALAEQKRAEEAAVAVMVASCPNHPRSVSTADNIYVNSHFEKSIKAGA 1320
Qy 1321 ISNPLPNRQF 1332
Db 1321 ISNPLPNRQF 1332

RESULT 5
AAG65638
ID AAG65638 standard; protein; 1332 AA.
XX
AC AAG65638;
XX
DT 07-JAN-2002 (first entry)
XX
DE SSD-containing SSP1 protein sequence.

XX SSD; sterol-sensing domain; human; liver; testis; brain; cancer;
KW nootropic; neuroprotective; antidiabetic; antiarteriosclerotic;
KW cytostatic; antilipemic; SSP1.
XX Homo sapiens.
XX WO200170974-A1.
XX 27-SEP-2001.
XX 22-MAR-2001; 2001WO-JP002279.
XX 24-MAR-2000; 2000JP-00088595.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Taniyama Y, Kita S, Komiyama T;
XX WPI; 2001-611501/70.
XX N-PSDB; RAI66914.
XX New sterol-sensing domain-containing protein for diagnosing and screening
PT candidate compounds in drug development for diabetes, obesity, cancer,
PT arteriosclerosis, hyperlipidemia and neurodegenerative disorders.
XX Claim 7; Page 115-122; 171pp; Japanese.
XX The invention provides a novel SSD (sterol-sensing domain)-containing
CC protein. The protein originates from human liver, human testis or human
CC brain. The protein can be expressed by standard recombinant methodology.
CC The proteins, encoded DNAs and antibodies are useful in diagnosis and
CC screening candidate compounds in drug development for diabetes, obesity,
CC cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such
CC as Alzheimer's disease and neural disorders. The present sequence
CC represents the SSP1 protein which contains the SSD domain
XX Sequence 1332 AA;
Qy Query Match 99.8%; Score 6896; DB 4; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60
Qy 61 TPARKITGDHLILLOKICPRLYTGNTOACCSAKOLVSLSEASITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGNTOACCSAKOLVSLSEASITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAPQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAPQSYDSCSRVRPAA 180
Qy 181 ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQLNEGV 240
Qy 241 ARCNEQGDVATCSCQDCAASCAPARPOALDSTFYLQMPGSLVLIILCSVFPAVTI 300
Db 241 ARCNEQGDVATCSCQDCAASCAPARPOALDSTFYLQMPGSLVLIILCSVFPAVTI 300
Qy 301 LLVGRVAPARKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTFELTDPVELMWSAPNSQARSKAFHQHFGFPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTFELTDPVELMWSAPNSQARSKAFHQHFGFPFRTNQVILTAPNRSY 420
Qy 421 RYDSELLGPKNFSGTLDLLELLELQERLRLHQLVMSPEAQRLNISLQDICYAPLNDPT 480

Db	421	RYDLSLLGPKNPSGILDLDLLELELEQLERLHLQVNSPEAQRNLSLQDICVAPLNPNT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL	540
Qy	541	SCMADYGAPVPPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKWEAFLEEM	600
Db	541	SCMADYGAPVPPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKWEAFLEEM	600
Qy	601	RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPATSYIVIFLYISIALGSSYSWSRV	660
Db	601	RAFQRMAGMFOVTTAERSLEDEINRTTAEDLPATSYIVIFLYISIALGSSYSWSRV	660
Qy	661	MVDSKATLGLGVAVVLGAVNAAMGFFSYLGIRSSLVILQVPPFLVLSVGADNIFVLE	720
Db	661	MVDSKATLGLGVAVVLGAVNAAMGFFSYLGIRSSLVILQVPPFLVLSVGADNIFVLE	720
Qy	721	YQRLPRRGPBREVHIGRALGRVAPSMLLCSLSEACFFLGNALTPMPAVRTFALTSLGLAV	780
Db	721	YQRLPRRGPBREVHIGRALGRVAPSMLLCSLSEACFFLGNALTPMPAVRTFALTSLGLAV	780
Qy	781	ILDFFLQMSAFVALLSLDSKQEAERLDVCCCVKPOELPPPGQGGGLLIGFQKAYAPPL	840
Db	781	ILDFFLQMSAFVALLSLDSKQEAERLDVCCCVKPOELPPPGQGGGLLIGFQKAYAPPL	840
Qy	841	LHWITRGVVLFLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVVLFLFALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Qy	901	YFVTLGNFSSSEAGMNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDDFDIM	960
Db	901	YFVTLGNFSSSEAGMNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDDFDIM	960
Qy	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSVQFHKYLPFLNDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCSTVNSLNCNCHSITWGSVRPSVQFHKYLPFLNDRP	1020
Qy	1021	NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL	1080
Db	1021	NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL	1080
Qy	1081	RKVPGTDPAPFVFPYTIINVFYEQVLTILPGLFMLSICLVPTFAVSCLLILGLDLRSGLL	1140
Db	1081	RKVPGTDPAPFVFPYTIINVFYEQVLTILPGLFMLSICLVPTFAVSCLLILGLDLRSGLL	1140
Qy	1141	NLLSTVMILVDVTGFMALWDISYNAVSLINLVSAVGMVSVEFVSHITRSFAISTKPTWLER	1200
Db	1141	NLLSTVMILVDVTGFMALWDISYNAVSLINLVSAVGMVSVEFVSHITRSFAISTKPTWLER	1200
Qy	1201	AKAATISMGSAVFAGVAMTNLPGIILVLGLAKAQLIQIFPFRINLLITLGLLHGLVFLPV	1260
Db	1201	AKAATISMGSAVFAGVAMTNLPGIILVLGLAKAQLIQIFPFRINLLITLGLLHGLVFLPV	1260
Qy	1261	ILSVGPDVNPALALEQKRAEEAAMVAVSCPNHPSRVSTADNIYNVHNSPEGSIKGAGA	1320
Db	1261	ILSVGPDVNPALALEQKRAEEAAMVAVSCPNHPSRVSTADNIYNVHNSPEGSIKGAGA	1320
Qy	1321	ISNLFNNGROF 1332	
Db	1321	ISNLFNNGROF 1332	
RESULT 6			
AA	AA	AA	AA
AC	AC	AC	AC
XX	XX	XX	XX
DT	DT	DT	DT
XX	XX	XX	XX
DE	DE	DE	DE
XX	XX	XX	XX

KW	Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation.
OS	Homo sapiens.
XX	WO200157190-A2.
XX	09-AUG-2001.
XX	05-FEB-2001; 2001WO-US004098.
XX	03-FEB-2000; 2000US-00496914.
XX	27-APR-2000; 2000US-00560875.
XX	20-JUN-2000; 2000US-00598075.
XX	19-JUL-2000; 2000US-00620325.
XX	01-SEP-2000; 2000US-00654936.
XX	15-SEP-2000; 2000US-00663561.
XX	20-OCT-2000; 2000US-00693325.
XX	30-NOV-2000; 2000US-00728422.
XX	(HYSE-) HYSEQ INC.
XX	Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y; Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW; Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX	WPI; 2001-476283/51.
XX	N-PSDB; AAK52301.
XX	Nucleic acids encoding polypeptides with cytokine-like activities, useful in diagnosis and gene therapy.
XX	Claim 20; Page 4213-4216; 6221pp; English.
XX	The invention relates to polynucleotides (AAK51456-AAK53435) and the encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to cytokine, cell proliferation or cell differentiation which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the sequence listing were missing at the time of publication
XX	Sequence 1359 AA;
Qy	Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Db	Best Local Similarity 97.9%; Pred. No. 0;
XX	Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
Qy	1 MAEAGLKGWLLWALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
Db	1 MAEAGLKGWLLWALLRLAQSEPTTTHQPGYCAFYDECCKNPELSGLMTLSNVCSLSN 60
Qy	61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db	61 TPARKITGDHLLILQKICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy	121 VNLCHTCSNPQSLFNTVTRVAQGLPAAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db	121 VNLCHTCSNPQSLFNTVTRVAQGLPAAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy	181 ATLAGVTMGVYGSALCNAQRWLNFGDGTGNGLAPLDTTHLLRPGQAVGSGIQPLNKG 240
Db	181 ATLAGVTMGVYGSALCNAQRWLNFGDGTGNGLAPLDTTHLLRPGQAVGSGIQPLNKG 240
Qy	241 ARCNEQDDVATCSCQDCACPAIARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300

Db	241		ARCNESQDDVATCSCQDCAACSPAIAARPOALDSTFYLGQMPGSLVLIILICSVFAVVTI	300
Qy	301		LLVGFVAPARDKSMVDPKKGTSLSKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360
Db	301		LLVGFVAPARDKSMVDPKKGTSLSKLSFSHTLLGQFFQGWGTWASWPLTILVLSV	360
Qy	361		IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILTAPNRSSY	420
Db	361		IPVVALAAGLVFTLTTPDVELMSAPNSQARSEKAFHDQHFQPPFRTNQVILTAPNRSSY	420
Qy	421		RYDSLLGPKNFSGLDLDLLELEQLERLRLHQLVMSPEAQRNLSLQDICYAPLNPDNT	480
Db	421		RYDSLLGPKNFSGLDLDLLELEQLERLRLHQLVMSPEAQRNLSLQDICYAPLNPDNT	480
Qy	481		SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALAL	540
Db	481		SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALAL	540
Qy	541		SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWBEAFLEEM	600
Db	541		SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWBEAFLEEM	600
Qy	601		RAFORMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSMSRV	660
Db	601		RAFORMAGMFQVTTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSMSRV	660
Qy	661		MYDSKATLGLGVAVVVLGAVNMAAGFFSYLGRSSLVILQVVPFLVLSVGADNFIPIFVLE	720
Db	661		MYDSKATLGLGVAVVVLGAVNMAAGFFSYLGRSSLVILQVVPFLVLSVGADNFIPIFVLE	720
Qy	721		YQRLPRRPGEPREHIGRAGRVAPSMLLCSLSEACFFLGLGALTMPMPAVRTFALTSGLAV	780
Db	721		YQRLPRRPGEPREHIGRAGRVAPSMLLCSLSEACFFLGLGALTMPMPAVRTFALTSGLAV	780
Qy	781		ILDFELLQMSAFVALLSLDSKREASRLDVCCKPQELPPCGQGLLGGFFQKAYAPFL	840
Db	781		ILDFELLQMSAFVALLSLDSKREASRLDVCCKPQELPPCGQGLLGGFFQKAYAPFL	840
Qy	841		LHWITRGVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLFNRYFVGVAPV	900
Db	841		LHWITRGVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLFNRYFVGVAPV	900
Qy	901		YFVTTLGYNFSFAGNNAICSSAGCNNFSTQKIQVATEFPFQSYLAIPASSWVDVDFDW	960
Db	901		YFVTTLGYNFSFAGNNAICSSAGCNNFSTQKIQVATEFPFQSYLAIPASSWVDVDFDW	960
Qy	961		LTPSSCCRLYISGPNKDKCPSVNSLNCIKNCMSITMGSVRPSVQPHKYLIPWFLNDRP	1020
Db	961		LTPSSCCRLYISGPNKDKCPSVNSLNCIKNCMSITMGSVRPSVQPHKYLIPWFLNDRP	1020
Qy	1021		NIKCPKGLAAYSTSVNLTSDGOVL-----ASRPMAYH 1053	
Db	1021		NIKCPKGLAAYSTSVNLTSDGOVLDTVALTSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080	
Qy	1054		KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL	1113
Db	1081		KPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL	1140
Qy	1114		FMLSCLVPTFAYSCILLGLDRLSGLNLASIVMILVDTVGFMAWDISVNAVSLNLVLS	1173
Db	1141		FMLSCLVPTFAYSCILLGLDRLSGLNLASIVMILVDTVGFMAWDISVNAVSLNLVLS	1200
Qy	1174		AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQ	1233
Db	1201		AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQ	1260
Qy	1234		LIQIFFPRLNLLITLGLLHGLVPLPVILSVGPDVNPALALBQKRAEEAAVAVMVASCP	1293
Db	1261		LIQIFFPRLNLLITLGLLHGLVPLPVILSVGPDVNPALALBQKRAEEAAVAVMVASCP	1320
Qy	1294		NHPSRVSTADNIYVNHSPFSGIKGAGISNPLPNNGRQF 1332	

Db 1321 NHPSRVSTADNIYVNHSPFSGIKGAGISNPLPNNGRQF 1359

RESULT 7

ADJ27337

ID ADJ27337 standard; protein; 1359 AA.

XX

AC ADJ27337;

XX

DT 20-MAY-2004 (first entry)

XX

DE Human NPC1L1.

XX

KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network; plasma membrane; transport signal; promoter;

KW sterol regulated element binding protein 1; SREBP1;

KW binding consensus sequence; transmembrane domain; sterol-sensing domain;

KW SCD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;

KW cholesterol absorption; serum cholesterol; hyperlipidaemia;

KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX

OS Homo sapiens.

XX

XX WO2004009772-A2.

XX

PD 29-JAN-2004.

XX

PF 17-JUL-2003; 2003WO-US022467.

XX

PR 19-JUL-2002; 2002US-0397442P.

XX

XX (SCHE) SCHERING CORP.

PA Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;

XX

PI WPI; 2004-132945/13.

DR N-PSDB; ADJ27336.

DR GENBANK; AF192522.

XX

PT New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting or identifying antagonists of NPC1L1 for inhibiting intestinal cholesterol absorption in a subject, or for treating elevated serum cholesterol or stroke.

XX

PS Example 7; SEQ ID NO 44; 125pp; English.

XX

CC This sequence represents a Niemann-Pick disease, type C1, gene-like 1 (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains a motif ADJ27331 which acts as a trans-golgi network to plasma membrane transport signal, and which exhibits limited tissue distribution and gastrointestinal abundance. The human NPC1L1 promoter sequence contains a sterol regulated element binding protein 1 (SREBP1) binding consensus sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-sensing domain (SSD) which is involved in sensing cholesterol levels, possibly by a mechanism which involves direct cholesterol bonding. NPC1L1 has 42% amino acid sequence homology to human NPC1, a receptor responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides of polynucleotides are useful for detecting or identifying antagonists of NPC1L1, which can be used for inhibiting intestinal cholesterol absorption in a subject, or for treating medical conditions including elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary heart disease, stroke or arteriosclerosis.

XX

SQ Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 8; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLUGMLLWALLRLAQSEPTTIHQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60

Db 1 MAEAGLUGMLLWALLRLAQSEPTTIHQPGYCAFYDECGKNPELSGLMTLSNVSLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNOSLEFINTVTRVAQLGAGOLPAVVAEAFVQHSFASQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNOSLEFINTVTRVAQLGAGOLPAVVAEAFVQHSFASQSDSCSRVRPAA 180
QY 181 ATLAUVTMCGVGSALCNARWLNFGQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
DB 181 ATLAUVTMCGVGSALCNARWLNFGQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
QY 241 ARCNEGQDDVATCSQDCAASCPAIPALQALDSTFYLQMPGSLVLIILILCSVPFAVVTI 300
DB 241 ARCNEGQDDVATCSQDCAASCPAIPALQALDSTFYLQMPGSLVLIILILCSVPFAVVTI 300
QY 301 LLAGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTVWASWPLTILVLSV 360
DB 301 LLAGFVAPARDKSKMVDPKGTSLSDKLSFSFTHLLGQFFQGWGTVWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
QY 421 RYDSLLGPNFSGIILDLLELLELQRLRLHQLWSPEAQRNLSLDICVAPLNPDT 480
DB 421 RYDSLLGPNFSGIILDLLELLELQRLRLHQLWSPEAQRNLSLDICVAPLNPDT 480
QY 481 SYDDCINSLLQYFQNNRTLLLTANQTLWGQTSQVDKDHFLYCANAPLTKQGTALAL 540
DB 481 SYDDCINSLLQYFQNNRTLLLTANQTLWGQTSQVDKDHFLYCANAPLTKQGTALAL 540
QY 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLPM 600
DB 541 SCHADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLPM 600
QY 601 RAFORMAGMFWTTFAERSLDEINRTTAEDLPATSYIIVIFLYISALGYSYSWSRV 660
DB 601 RAFORMAGMFWTTFAERSLDEINRTTAEDLPATSYIIVIFLYISALGYSYSWSRV 660
QY 661 MVDKATILGGVAVVILGAVMAAMGFFSYLGRIRSSLVILQVVPFLVLSVGAQNIFFVLE 720
DB 661 MVDKATILGGVAVVILGAVMAAMGFFSYLGRIRSSLVILQVVPFLVLSVGAQNIFFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALICFELGALTMPAVRTFALTSGLA 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALICFELGALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKPQELPPQGGEGLLGFPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCKPQELPPQGGEGLLGFPQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFVGARV 900
DB 841 LHWITRGVLLFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFLNRYFVGARV 900
QY 901 YFVTTLGYNFSZAGNNAICSSAGCNFSTQKIOVATEPFPQSQYLAIIPASSWVDPDFID 960
DB 901 YFVTTLGYNFSZAGNNAICSSAGCNFSTQKIOVATEPFPQSQYLAIIPASSWVDPDFID 960
QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMCSITMGSVRPSVEQFHKYLPWFINDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMCSITMGSVRPSVEQFHKYLPWFINDRP 1020
QY 1021 NTKCPKGGLAAYSTSNLTSQDVL-----ASRFMAYH 1053
DB 1021 NTKCPKGGLAAYSTSNLTSQDVL-----ASRFMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEFVPYTTINVFYEQVLTILPEGL 1113
DB 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEFVPYTTINVFYEQVLTILPEGL 1113
QY 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEFVPYTTINVFYEQVLTILPEGL 1140
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDAPEFVPYTTINVFYEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLNLLSIIVMLVDTVGFPMALWDISYNAVSLINLVS 1173

DB 1141 FMLSCLVPTFAVSCLLGLDLRLSGLNLLSIIVMLVDTVGFPMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATTISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFFRLNLLITLGLLHGLVFLFPVILSYVYVDPVNPALAEQKRAEBAVAAMVVASCP 1293
DB 1261 LIQIPFFRLNLLITLGLLHGLVFLFPVILSYVYVDPVNPALAEQKRAEBAVAAMVVASCP 1320
QY 1294 NHPRSVSTADNIIYVNHSPFSGIKGAGAIISNPLPNNGRQF 1332
DB 1321 NHPRSVSTADNIIYVNHSPFSGIKGAGAIISNPLPNNGRQF 1359
RESULT 8
ADY60780
ID ADY60780 standard; protein; 1359 AA.
XX AC ADY60780;
XX DT 19-MAY-2005 (first entry)
XX DE Human Niemann-Pick C1-like protein 1 (NPC1L1) protein SEQ ID 44.
XX KW Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
KW Cardiant; Vasotropic; Cerebroprotective; cholesterol; hyperlipidemia;
KW metabolic disorder; atherosclerosis; cardiovascular disease;
KW coronary artery disease; cerebrovascular ischemia; neurological disease;
KW arteriosclerosis; transgenic animal.
XX OS Homo sapiens.
XX PN WO2005015988-A1.
XX PD 24-FEB-2005.
XX PF 16-DEC-2003; 2003WO-US040113.
XX PR 17-JUL-2003; 2003US-00621758.
PR 22-AUG-2003; 2003US-00646301.
PR 16-SEP-2003; 2003US-00663208.
XX (SCHE) SCHERING CORP.
XX PI Altman SW, Murgolo NJ, Wang LQ, Graziano MP;
XX WPI; 2005-284403/29.
XX N-PSDB; ADY60779.
PT Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
PT NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
PT treating hyperlipidemia or atherosclerosis in mouse, rat or human.
PS Disclosure; SEQ ID NO 44; 146pp; English.
XX The invention relates to an isolated Niemann-Pick C1-like protein 1
CC (NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
CC amino acid sequence appearing as ADY60738,ADY60740 or ADY60748 (Rat,
CC human and Mouse NPC1L1). Also included are an isolated polynucleotide
CC encoding NPC1L1 (appearing as ADY60737,ADY60739 or ADY60747), a
CC recombinant vector comprising the polynucleotide, a host cell comprising
CC the vector, an isolated antibody which specifically binds to NPC1L1 (or
CC to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
CC preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
CC mutation of endogenous chromosomal NPC1L1 (where the mouse does not
CC produce any functional NPC1L1 protein), an offspring or progeny of the
CC mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
CC sample for an intestinal sterol or Salpha-sterol absorption antagonist, a
CC inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject, a
CC kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
CC dosage form, and information indicating that NPC1L1 is a target of

ezetimibe or substituted azetidinone), decreasing the level of intestinal sterol or Salpha-stanol absorption in a subject (involving reducing the level of expression of NPC11l1 in the subject), identifying an antagonist of NPC11l1 and an isolated mammalian cell which lacks a gene encoding a functional NPC11l1 protein. NPC11l1 is useful for identifying an antagonist of NPC11l1 which is useful for inhibiting or decreasing the level of NPC11l1 mediated sterol or Salpha-stanol uptake in a subject such as mouse, rat or human. The level of expression of NPC11l1 in the subject is reduced by mutating NPC11l1 in the subject. The NPC11l1 knock-out mouse is useful for screening a sample for intestinal sterol or Salpha-stanol absorption antagonist. The NPC11l1 antagonists identified are useful for treating medical conditions caused or mediated by NPC11l1, e.g., hyperlipidemia, atherosclerosis, coronary heart diseases, stroke or arteriosclerosis. The present sequence represents a mammalian NPC11l1 protein.

XX	Sequence 1359 AA;	Query Match Best Local Similarity Matches 1330; Conservative	99.5%; 97.9%; 0;	Score 6872.5; Pred. No. 0; 0;	DB 9;	Length 1359;	27;	Gaps 1;
QY	1	MAEAGLRGMLLWALLRLAQSEBYTTIHQBGYCAFYDECCKNPDELGSMTLSNVCSLSN	60					
Db	1	MAEAGLRGMLLWALLRLAQSEBYTTIHQBGYCAFYDECCKNPDELGSMTLSNVCSLSN	60					
QY	61	TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF	120					
Db	61	TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF	120					
QY	121	VNLHCHNTCPNQSLFINVTRVLAQLGAGQLPAVVAVYEAFYQHSFAEQSYDSCSRVRVPAA	180					
Db	121	VNLHCHNTCPNQSLFINVTRVLAQLGAGQLPAVVAVYEAFYQHSFAEQSYDSCSRVRVPAA	180					
QY	181	ATLAVGTMCGVYSALCNAORWLNFGDTCNGGLAPLDITPHLLRPGQAVSGGIQPLNEGV	240					
Db	181	ATLAVGTMCGVYSALCNAORWLNFGDTCNGGLAPLDITPHLLRPGQAVSGGIQPLNEGV	240					
QY	241	ARCNESQGDVDVATCSQDCAASCPIAIPALQALDSTFYLGQMPGSLVLIILCSVFVAVVTI	300					
Db	241	ARCNESQGDVDVATCSQDCAASCPIAIPALQALDSTFYLGQMPGSLVLIILCSVFVAVVTI	300					
QY	301	LLVGFRVAPARDKSKMVDPKKGTSLSKLSFTSHITLGGPFGQGWTVWASWPLTILVLSV	360					
Db	301	LLVGFRVAPARDKSKMVDPKKGTSLSKLSFTSHITLGGPFGQGWTVWASWPLTILVLSV	360					
QY	361	IPVVALAAGLVFTLETTDPVELMSAPNSQARSEKAFHDQHGPFRTNQVILTAPNRSSY	420					
Db	361	IPVVALAAGLVFTLETTDPVELMSAPNSQARSEKAFHDQHGPFRTNQVILTAPNRSSY	420					
QY	421	RYDSLLLGPKNFSIGILDLLLELELEQERLRLHQWSPEAQRNISIQDICYAPLNPDNT	480					
Db	421	RYDSLLLGPKNFSIGILDLLLELELEQERLRLHQWSPEAQRNISIQDICYAPLNPDNT	480					
QY	481	SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGOTALAL	540					
Db	481	SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVDWKDHFLYCANAPLTFKDGOTALAL	540					
QY	541	SCMADYGAPVFPFLAIGYKGKDYSEALIMTFTSLNNYPAGDPRPLAQAKLWEEAFLEEM	600					
Db	541	SCMADYGAPVFPFLAIGYKGKDYSEALIMTFTSLNNYPAGDPRPLAQAKLWEEAFLEEM	600					
QY	601	RAFQRRMAGMPOVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSYSSWRV	660					
Db	601	RAFQRRMAGMPOVFTFAERSLEDEINRTTAEDLPIFATSYIVIFLYISIALGYSYSSWRV	660					
QY	661	MVDSKATILGLGGVAVVLGAVMAAMGFFSYGLGTRSSLVILQVVPFLVLSVGADNIPFVLE	720					
Db	661	MVDSKATILGLGGVAVVLGAVMAAMGFFSYGLGTRSSLVILQVVPFLVLSVGADNIPFVLE	720					
QY	721	YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPAVRFTALTSGLAV	780					
Db	721	YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFFLGLALTMPAVRFTALTSGLAV	780					

QY	781	ILDFLQMSAFVALLSLDSKRQBSRLDYCCCVKQELPPGQEGILLGFFQKAYAPFL	840
DB	781	ILDFLQMSAFVALLSLDSKRQBSRLDYCCCVKQELPPGQEGILLGFFQKAYAPFL	840
QY	841	LHWITRGVLLLFALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	900
DB	841	LHWITRGVLLLFALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA	900
QY	901	YFVTLLGYNFSFAGHNAICSSAGCNFSTQKIYATEPPEQSYLAIPASSWDDFD	960
DB	901	YFVTLLGYNFSFAGHNAICSSAGCNFSTQKIYATEPPEQSYLAIPASSWDDFD	960
QY	961	LTSSCCRLYISGPNKDKFCPSTVNSLNCUKNCISITWGSVRPSVEQFHKYL	1020
DB	961	LTSSCCRLYISGPNKDKFCPSTVNSLNCUKNCISITWGSVRPSVEQFHKYL	1020
QY	1021	NIKCPKGGLAAYSTSVNLTSDGOVL-----ASRFMAYH	1053
DB	1021	NIKCPKGGLAAYSTSVNLTSDGOVL-----ASRFMAYH	1053
QY	1054	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTNVFYEQY	1113
DB	1081	KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTNVFYEQY	1140
QY	1114	FMLSLCLVPTFAVSCLLGLDLRSGLANLSIVMLVDTVGFMALWDISYNAVS	1173
DB	1141	FMLSLCLVPTFAVSCLLGLDLRSGLANLSIVMLVDTVGFMALWDISYNAVS	1200
QY	1174	AVGMSVEFVSHITRSPAIKPTWLERAKEATISMGSAVFAGVAMTNLPGIL	1233
DB	1201	AVGMSVEFVSHITRSPAIKPTWLERAKEATISMGSAVFAGVAMTNLPGIL	1260
QY	1234	LIQIFFFRLNLLITLLGLHLVFLPVILSYGPDVNPALALEQKRAEAAVAV	1293
DB	1261	LIQIFFFRLNLLITLLGLHLVFLPVILSYGPDVNPALALEQKRAEAAVAV	1320
QY	1294	NHPSRVSTADNIYVNHSPFEGSIKAGAIISNPLPNNRGF	1332
DB	1321	NHPSRVSTADNIYVNHSPFEGSIKAGAIISNPLPNNRGF	1359
RESULT 9			
AEB93610			
XX	ID	AEB93610 standard; protein; 1359 AA.	
XX	AC	AEB93610;	
XX	DT	06-OCT-2005 (first entry)	
XX	DE	Human NPC1L1 protein.	
XX	KW	Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;	
XX	KW	cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;	
XX	KW	arteriosclerosis; ds.	
XX	OS	Homo sapiens.	
XX	PN	WO2005069900-A2.	
XX	PD	04-AUG-2005.	
XX	PF	14-JAN-2005; 2005WO-US001469.	
XX	PR	16-JAN-2004; 2004US-0537341P.	
XX	PA	(MERI) MERCK & CO INC.	
XX	PI	Garcia-Calvo M;	
XX	DR	WPI; 2005-564070/57.	
XX	DR	N-PSDB; AEB93609.	
XX	DR	GENBANK; AF192522.	

XX Identifying ligand of NPC111 for stimulating the activity of NPC111, by
PT contacting NPC111 with detectably labeled substituted 2-azetidinone
PT glucuronide and a candidate compound and determining if compound binds to
PT human NPC111.

Disclosure; SEQ ID NO 44; 215pp; English.

The invention relates to identifying a ligand of NPC11L1. The method involves contacting human NPC11L1 with a detectably labeled substituted 2-azetidinone glucuronide and a candidate compound and determining whether the candidate compound binds to human NPC11L1. In identifying a ligand of NPC11L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10 nM or lower. The detectably labeled substituted 2-azetidinone glucuronide is 35 S-labeled compound 2, given in the specification. NPC11L1 ligands are useful for stimulating or blocking the activity of NPC11L1, and for treating conditions caused or mediated by NPC11L1. It is useful for reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart disease, stroke, or arteriosclerosis. The present sequence represents the amino acid sequence of human NPC11L1 as disclosed under GenBank accession number AF192522.

Sequence 1359 AA;

Query Match 99.5%; Score 6872.5; DB 9; Length 1359;

Best Local Similarity 97.9%; Pred. No. 0;

Accession	Length	Matches	Conservative	Mismatches	Indels	Gaps
U00000	1330	1330	0	0	2	27
U00001	1330	1330	0	0	2	27
U00002	1330	1330	0	0	2	27
U00003	1330	1330	0	0	2	27
U00004	1330	1330	0	0	2	27
U00005	1330	1330	0	0	2	27
U00006	1330	1330	0	0	2	27
U00007	1330	1330	0	0	2	27
U00008	1330	1330	0	0	2	27
U00009	1330	1330	0	0	2	27
U00010	1330	1330	0	0	2	27
U00011	1330	1330	0	0	2	27
U00012	1330	1330	0	0	2	27
U00013	1330	1330	0	0	2	27
U00014	1330	1330	0	0	2	27
U00015	1330	1330	0	0	2	27
U00016	1330	1330	0	0	2	27
U00017	1330	1330	0	0	2	27
U00018	1330	1330	0	0	2	27
U00019	1330	1330	0	0	2	27
U00020	1330	1330	0	0	2	27
U00021	1330	1330	0	0	2	27
U00022	1330	1330	0	0	2	27
U00023	1330	1330	0	0	2	27
U00024	1330	1330	0	0	2	27
U00025	1330	1330	0	0	2	27
U00026	1330	1330	0	0	2	27
U00027	1330	1330	0	0	2	27
U00028	1330	1330	0	0	2	27
U00029	1330	1330	0	0	2	27
U00030	1330	1330	0	0	2	27
U00031	1330	1330	0	0	2	27
U00032	1330	1330	0	0	2	27
U00033	1330	1330	0	0	2	27
U00034	1330	1330	0	0	2	27
U00035	1330	1330	0	0	2	27
U00036	1330	1330	0	0	2	27
U00037	1330	1330	0	0	2	27
U00038	1330	1330	0	0	2	27
U00039	1330	1330	0	0	2	27
U00040	1330	1330	0	0	2	27
U00041	1330	1330	0	0	2	27
U00042	1330	1330	0	0	2	27
U00043	1330	1330	0	0	2	27
U00044	1330	1330	0	0	2	27
U00045	1330	1330	0	0	2	27
U00046	1330	1330	0	0	2	27
U00047	1330	1330	0	0	2	27
U00048	1330	1330	0	0	2	27
U00049	1330	1330	0	0	2	27
U00050	1330	1330	0	0	2	27
U00051	1330	1330	0	0	2	27
U00052	1330	1330	0	0	2	27
U00053	1330	1330	0	0	2	27
U00054	1330	1330	0	0	2	27
U00055	1330	1330	0	0	2	27
U00056	1330	1330	0	0	2	27
U00057	1330	1330	0	0	2	27
U00058	1330	1330	0	0	2	27
U00						

Qy	1	MAEAGLRGWLWALLRLRLAQSPEYTTIHQPGYCAPYDECGKRNPELSGSLMTLSNVCSLSN	60
Db	1	MAEAGLRGWLWALLRLRLAQSPEYTTIHQPGYCAPYDECGKRNPELSGSLMTLSNVCSLSN	60
Qy	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSEASLSITTKALLTRCPACSNF	120
Db	61	TPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSEASLSITTKALLTRCPACSNF	120
Qy	121	VNLCHNTCSNQSLFINVTRVAOLGAGOLPAVVAEYAFYQHSFAEQSYDSCSRVRVPAA	180
Db	121	VNLCHNTCSNQSLFINVTRVAOLGAGOLPAVVAEYAFYQHSFAEQSYDSCSRVRVPAA	180
Qy	181	ATLAVGTMCVGYGSALCNAQRWLNFQDGTGNGLAPLDITFHLLEPGQAVGSGIQLPNSGV	240
Db	181	ATLAVGTMCVGYGSALCNAQRWLNFQDGTGNGLAPLDITFHLLEPGQAVGSGIQLPNSGV	240
Qy	241	ARCNESQGDVVATCSQDCAASCFAIARPAQALDSTFYLGQMPGSLVLIILCSFEAVVTI	300
Db	241	ARCNESQGDVVATCSQDCAASCFAIARPAQALDSTFYLGQMPGSLVLIILCSFEAVVTI	300
Qy	301	LLVGFVPAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFQGGCTWVASWPLTITLVLSV	360
Db	301	LLVGFVPAPARDKSNVDPKKGTSLSDKLSFSTHTLLGQFQGGCTWVASWPLTITLVLSV	360
Qy	361	IPVVAALAGLVFTBELTTPDVELWAPNSQARSEKAFHQHGGPPFRTNQVILTAENRSSY	420
Db	361	IPVVAALAGLVFTBELTTPDVELWAPNSQARSEKAFHQHGGPPFRTNQVILTAENRSSY	420
Qy	421	RYDSLILGPKNFSGILDLDLLELLELQERLRLHVQWSPERQARNISLODICIYAPLNPNPT	480
Db	421	RYDSLILGPKNFSGILDLDLLELLELQERLRLHVQWSPERQARNISLODICIYAPLNPNPT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALAL	540
Qy	541	SCMADYGAPVPPFLAIGYKGKQYSEBALIMTFPSLNNYPAGDPRPRLAQAKLWEERAFLEEM	600
Db	541	SCMADYGAPVPPFLAIGYKGKQYSEBALIMTFPSLNNYPAGDPRPRLAQAKLWEERAFLEEM	600
Qy	601	RAFORRWAGMEQVTFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV	660
Db	601	RAFORRWAGMEQVTFTAERSLEDEINRTTAEDLPIFATSYIVIFLYISLAGSYSSWSRV	660
Qy	661	MVDSKATILGLGVAWLVGAVNMAAGFYSYLGISSVLIVQVVPFLVLSVGADNIPFIIVLE	720

RESULT 10

RESULT IN
ABG22693

ABG22693
ID ABG22693 standard: protein: 1344 AA.

XX ABG22693:

AC XX
ABG22693;

DT 18-FEB-2002 (first entry)

18-FEB-2002 (11:50 AM)
XX
DE Novel human diagnostic protein #22684.

XX
DE NOVEMBER 1951

KW Human; chromosome

human, children
food supplement

[illegible]

OS Homo sapiens.

XX
name of the

PN WO200175067-A2.

XX	Drmanac RT, Liu C, Tang YT;	QY	479	NTSLYDCCINSLLQYFQNNRTLLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTAL	538
PI		Db	479	NTSLYDCCINSLLQYFQNNRTLLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTAL	538
DR	WPI; 2001-639362/73.	QY	539	ALSCMADYGAIPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFL	598
DR	N-PSDB; AAS86880.	Db	539	ALSCMADYGAIPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFL	598
XX	New isolated polynucleotide and encoded polypeptides, useful in	QY	599	EMRAFORRMAGMFOVTFPAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSYSWS	658
PT	diagnostics, forensics, gene mapping, identification of mutations	Db	599	EMRAFORRMAGMFOVTFPAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSYSWS	658
PT	responsible for genetic disorders or other traits and to assess	QY	659	RVWVDSKATLGLGGAIVVLGAVMAAMGFFSYLGRSSLVILQVVPFVLVSGADNIFIV	718
PT	biodiversity.	Db	659	RVWVDSKATLGLGGAIVVLGAVMAAMGFFSYLGRSSLVILQVVPFVLVSGADNIFIV	718
PS	Claim 20; SEQ ID NO 53052; 103pp; English.	QY	719	LEYQRLPRRGPGEPRVHIGRALGRVAPSMILCSLSEACFPLGALTMPAVRTALTSGL	778
XX	The invention relates to isolated polynucleotide (I) and polypeptide (II)	Db	719	LEYQRLPRRGPGEPRVHIGRALGRVAPSMILCSLSEACFPLGALTMPAVRTALTSGL	778
CC	sequences. (I) is useful as hybridisation probes, polymerase chain	QY	779	AVILDFLLQMSAFVALLSLDSKROEASRLDVCCVKQELPPPGQGGELLGFFQKAYAP	838
CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,	Db	779	AVILDFLLQMSAFVALLSLDSKROEASRLDVCCVKQELPPPGQGGELLGFFQKAYAP	838
CC	and in recombinant production of (II). The polynucleotides are also used	QY	839	FLHWHITRGVV---LLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF	894
CC	in diagnostics as expressed sequence tags for identifying expressed	Db	839	FLHWHITRGVV---LLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYF	894
CC	genes. (I) is useful in gene therapy techniques to restore normal	QY	895	EVGAPVPTVTLGVNFSSEAGMNAICSSAGCNSFTQKIQYATEFPEQSYLSUPASSWV	954
CC	activity of (II) or to treat disease states involving (II). (II) is	Db	895	EVGAPVPTVTLGVNFSSEAGMNAICSSAGCNSFTQKIQYATEFPEQSYLSUPASSWV	954
CC	useful for generating antibodies against it, detecting or quantitating a	QY	955	DDFDMLTPSCCRLYISGPNKDFCSTVNSLNCNCSITMGSVRPSVEQFHKYLWP	1014
CC	polypeptide in tissue, as molecular weight markers and as a food	Db	955	DDFDMLTPSCCRLYISGPNKDFCSTVNSLNCNCSITMGSVRPSVEQFHKYLWP	1014
CC	supplement. (II) and its binding partners are useful in medical imaging	QY	1015	FLNDRPNKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAA	1074
CC	of sites expressing (II). (I) and (II) are useful for treating disorders	Db	1015	FLNDRPNKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAA	1074
CC	involving aberrant protein expression or biological activity. The	QY	1075	NITADLRKVPDTPAFEPFPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLD	1134
CC	polypeptide and polynucleotide sequences have application in	Db	1075	NITADLRKVPDTPAFEPFPTITNVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLD	1134
CC	diagnostics, forensics, gene mapping, identification of mutations	QY	1135	LRSGLLNLLSIWMLVDTGVFMALWDISYNAVSLINLVSAGMSVEFVSHITRFAISTK	1194
CC	responsible for genetic disorders or other traits to assess biodiversity	Db	1135	LRSGLLNLLSIWMLVDTGVFMALWDISYNAVSLINLVSAGMSVEFVSHITRFAISTK	1194
CC	and to produce other types of data and products dependent on DNA and	QY	1199	PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRMLNLLITLGLLHG	1254
CC	amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic	Db	1199	PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRMLNLLITLGLLHG	1254
CC	patent did not appear in the printed specification, but was obtained in	QY	1255	LVFLPVILSYGVDPVNPALALAKAEAE-----VAAMVVASCPNHPSPVST	1301
CC	electronic format directly from WIPO at	Db	1255	LVFLPVILSYGVDPVNPALALAKAEAE-----VAAMVVASCPNHPSPVST	1301
CC	ftp.wipo.int/pub/published_pct_sequences	QY	1259	LVFLPVILSYGVDPVNPALAEQKRAEGGGSHGGLAQITPPSPQLTTSMT	1313
XX		Db	1259	LVFLPVILSYGVDPVNPALAEQKRAEGGGSHGGLAQITPPSPQLTTSMT	1313
SQ	Sequence 1344 AA;				
	Query Match 94.6%; Score 6536; DB 4; Length 1344;				
	Best Local Similarity 96.7%; Pred. No. 0;				
	Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;				
QY	1 MAEAGLGLWLLALLRLAQSEPYTHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60				
Db	1 MAEAGLGLWLLALLRLAQSEPYTHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60				
QY	61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVLSLEASISITKALLTRCPACSDNF 120				
Db	61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVLSLEASISITKALLTRCPACSDNF 120				
QY	121 VNLHCHNTCSPNQSLFINTRVAQLGAGQLPAVVAEYAFQHSFAPQSDSCSRVRPAA 180				
Db	121 VNLHCHNTCSPNQSLFINTRVAQLGAGQLPAVVAEYAFQHSFAPQSDSCSRVRPAA 180				
QY	181 ATLAVGTMCGVYSALCNAORWLNFGDGTGNGLAPLDDITPHLLEPGQAVSGIQPLNEGV 240				
Db	181 ATLAVGTMCGVYSALCNAORWLNFGDGTGNGLAPLDDITPHLLEPGQAVSGIQPLNEGV 240				
QY	241 ARCNEOGDDVATCSQDCQAACSP--AIAFPQALDSTFYLGQMPGSLVLILILCSVFVV 298				
Db	241 ARCNEOGDDVATCSQDCQAACSP--AIAFPQALDSTFYLGQMPGSLVLILILCSVFVV 298				
QY	299 TILLVGRVAPARDKSRWDPKKGTSLSDKLSFSTHTLLQPFQGGKGTWVASWPLTILVL 358				
Db	299 TILLVGRVAPARDKSRWDPKKGTSLSDKLSFSTHTLLQPFQGGKGTWVASWPLTILVL 358				
QY	359 SVIEPVALLAGLVTELTTPDVELWSAPNSQAEKAFHDOHFGPPPTNQVILITAPNRS 418				
Db	359 SVIEPVALLAGLVTELTTPDVELWSAPNSQAEKAFHDOHFGPPPTNQVILITAPNRS 418				
QY	419 SYRYDSLLGLPKNPSGILDLDLLLELLELQERLHLQWSPAEQRNLSLQDICVAPLNPD 478				
Db	419 SYRYDSLLGLPKNPSGILDLDLLLELLELQERLHLQWSPAEQRNLSLQDICVAPLNPD 478				

RESULT 11
ADJ27295
ID ADJ27295 standard; protein; 1331 AA.
XX
AC ADJ27295;
XX
DT 20-MAY-2004 (first entry)
XX
DE Rat NPC1L1.
XX
KW Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; protein 1; SREBP1;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;

ADY60738;
19-MAY-2005 (first entry)
Rat Niemann-Pick C1-like protein 1 (NPC1L1) protein.
Niemann-Pick C1-like protein 1; Antilipemic; Antiarteriosclerotic;
Cardiant; Vasotrophic; Cerebroprotective; cholesterol; hyperlipidemia;
Metabolic disorder; atherosclerosis; cardiovascular disease;
coronary artery disease; cerebrovascular ischemia; neurological disease;
arteriosclerosis; transgenic animal.
Rattus sp.
WO2005015988-A1.
24-FEB-2005.
16-DEC-2003; 2003WO-US040113.
17-JUL-2003; 2003US-00621758.
22-AUG-2003; 2003US-00646301.
16-SEP-2003; 2003US-00663208.
(SCHE) SCHERING CORP.
Altmann SW, Murgolo NJ, Wang LQ, Graziano MP;
WPI; 2005-284403/29.
N-PSDB; ADY60737, ADY60746.
Novel isolated NPC1L1 polypeptide, useful for identifying antagonist of
NPC1L1 that inhibits NPC1L1 mediated sterol or 5 alpha-sterol uptake, for
treating hyperlipidemia or atherosclerosis in mouse, rat or human.
Claim 30; SEQ ID NO 2; 146pp; English.
The invention relates to an isolated Niemann-Pick C1-like protein 1
(NPC1L1) polypeptide comprising 42 or more contiguous amino acids from an
amino acid sequence appearing as ADY60738, ADY60740 or ADY60748 (Rat,
human and mouse NPC1L1). Also included are an isolated polynucleotide
encoding NPC1L1 (appearing as ADY60737, ADY60739 or ADY60747), a
recombinant vector comprising the polynucleotide, a host cell comprising
the vector, an isolated antibody which specifically binds to NPC1L1 (or
to a rat NPC1L1 peptide sequence appearing as ADY60775-ADY60778),
preparation of NPC1L1, a mutant transgenic mouse comprising a homozygous
mutation of endogenous chromosomal NPC1L1 (where the mouse does not
produce any functional NPC1L1 protein), an offspring or progeny of the
mouse having inherited a mutated NPC1L1 allele of the mouse, screening a
sample for an intestinal sterol or Salpha-sterol absorption antagonist,
inhibiting NPC1L1 mediated sterol or Salpha-sterol uptake in a subject,
kit (comprising ezetimibe or substituted azetidinone in a pharmaceutical
dosage form, and information indicating that NPC1L1 is a target of
ezetimibe or substituted azetidinone), decreasing the level of intestinal
sterol or Salpha-sterol absorption in a subject (involving reducing the
level of expression of NPC1L1 in the subject), identifying an antagonist
of NPC1L1 and an isolated mammalian cell which lacks a gene encoding a
functional NPC1L1 protein. NPC1L1 is useful for identifying an antagonist
of NPC1L1 which is useful for inhibiting or decreasing the level of
NPC1L1 mediated sterol or Salpha-sterol uptake in a subject such as
mouse, rat or human. The level of expression of NPC1L1 in the subject is
reduced by mutating NPC1L1 in the subject. The NPC1L1 knockout mouse is
useful for screening a sample for intestinal sterol or Salpha-sterol
absorption antagonist. The NPC1L1 antagonists identified are useful for
'treating medical conditions caused or mediated by NPC1L1, e.g.,
hyperlipidemia, atherosclerosis, coronary heart disease, stroke or
arteriosclerosis. The present sequence represents a mammalian NPC1L1
protein.
Query Match 78.5%; Score 5421.5; DB 9; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;

		Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;			
QY	1	MAEAGLRGWLWALLRLA	SEPYTHIQPGYCAFYDE	GKPNELSGSLMTLSNV	CLSN 60
DB	1	MAAAWL-GWLLWALLSAA	OGELYTPKHEAGVCTT	FEECGKPNELSGSLT	SNVCLSN 59
QY	61	TPARKITGHLILLQIKI	CPRLYTGPNTO-ACCSA	KOLVSLSEASLSITK	ALLTRCPACSDN 119
DB	60	TPARHTVGEHALLORI	CPRLYNGPNYTFACST	KQLLSLESSMSITK	ALLTRCPACSDN 119
QY	120	FVNHLCHNTCSNQSLF	FINVTRVAQIAGAGOL	PAVAYEAPYQHSFAE	OSVDSRVRVPA 179
DB	120	FVSLCHNTCSPOSLF	FINVTRVVERGAGEPP	AVAYEAFYQSFPAEK	AYESCQVRIPA 179
QY	180	AATLAVTMCVGYGSA	LNAQRWLNFGDGTNG	LAPLDITFHLLPEQA	AVGSGIQPLNEG 239
DB	180	AASLAVGSMCGVYGS	ALCNAQRWLNFGDGT	NGLAPLDITFHLLPE	QALPDGIQPLNGK 239
QY	240	VARNESQGDVATCSQ	DCCAASCPAIRPAQAL	DSFTYLGQMPGSLV	LIILCSVPAVVT 299
DB	240	IAPNBSQGDSDAVCS	QDCCAASCPVIPPPE	ALRPSFYMGMPGWL	ALIIIFTAVFVLLS 299
QY	300	ILLVGFVAPARDKSM	VDPKKTSLSDKLSFT	HTILLGQFQCGTWA	SWPLTILVLS 359
DB	300	AVLVRLVWVSNRNK	KAEGPQEAAPKLPHK	JKLSPHTILGRFFQ	NWGRVVASWPLVLS 359
QY	360	VIPVVALAAGLVFTE	LTTPDVELWSAPNSO	ARSEKAFHQHFGFF	FTNVOVILTAPNRSS 419
DB	360	FIVVIALAAGLTFT	ELTTDPVELWSAPNS	QARSEKSFHDEHFG	FFFTNQIFVTARNSS 419
QY	420	YRYSLLGLPKNFSG	IILDLLELQERLRH	LQVMSPEAQNISLO	DIQCYAPLAPDN 479
DB	420	YKYSLLGLSKNFSG	IILDLLELQERLRH	LQVMSPEAQNISLO	DIQCYAPLAPDN 479
QY	480	TSLVDCINSLLOQF	QNNRTLLLTANQTL	MGQTSQVDMKDHLY	CANAPLTKDGTALA 539
DB	480	TSLSDDCVNSLLOQ	FQNNRTLLLTANQTL	NGQTSQVDMKDHLY	CANAPLTKDGTALA 539
QY	540	LSCWADYGA	VPFPLATGGYKGDY	SEAEALIMTFSLNNY	PAGDPRLAAKLWEEAFLEE 599
DB	540	LSCWADYGA	VPFPLATGGYKGDY	SEAEALIMTFSLNNY	PAGDPRLAAKLWEEAFLEE 599
QY	600	MRAFQRMAGMFQV	TFTAERSLEDEIN	RTTAEDLPIFATS	YIVIFLYISLALGYSVSSWR 659
DB	600	MESQRTNSDKFQV	AFSAERSLEDEIN	RTTQDLQVFAVS	YIIVFLYISLALGYSVSSWR 659
QY	660	VWVDSKATLG	GGVAVVLGAVMA	MGPSYLGISSVILQ	VWVPLVLSVGDNIFIFVL 719
DB	660	VAVESKATLG	GGVAVVLGAVMA	MGPSYLGISSVILQ	VWVPLVLSVGDNIFIFVL 719
QY	720	EYQRLPRGP	PREVHIGRALGR	VAPSMLLCSLSA	ICFFLGALTPMPAVRTFALTSGLA 779
DB	720	EYQRLPRGP	PREVHIGRALGR	VAPSMLLCSLSA	ICFFLGALTPMPAVRTFALTSGLA 779
QY	780	VILDFLLQMSA	FPVALLSLSQKQ	EAASRLDVCCCVK	PQELPPPGQEGILLGFQKAYAPP 839
DB	780	IILDFLLQMTA	FVALLSLSQKQ	EAASRLDVCCCVK	PQELPPPGQEGILLGFQKAYAPP 839
QY	840	LLHWITRGV	LLVLLFLALGVS	LSYNSCHISVGLDQ	BALPKDSYLLDYFLFNRYFEVGAP 899
DB	840	LLHREIRPV	WMLFLTLFGAN	LYLMCNINVGUDQ	ELALPKDSYLLDYFLFNRYLEVGP 899
QY	900	VYFVTTLV	GVNFSSEAGMNA	ICSSAGCNPFSTQ	KIQYATEPPEQSYLAIIPASSWDDFID 959
DB	900	VYFVTTLV	GVNFSSEAGMNA	ICSSAGCNPFSTQ	KIQYATEPPEQSYLAIIPASSWDDFID 959
QY	960	WLTP-SSCRL	YISGPNKDKFCP	STVNSLNCILKNC	MSITMGSVRPSVVEGFKYLPWFLND 1018
DB	960	WLTPSSCRL	YIRGPHKDEF	CPSTDSFNCLKNC	MNRTLGVPVPTAEQFHKYLPWFLND 1019
QY	1019	RPNIKCPK	GLAAYSTVNL	TSDDGVLASRF	WAYHKLKNSODYTEALRAARELAANITA 1078
DB	1020	PNNIRCPK	GLAAIRTSVNL	SSDGGQVIA	SOFMAYHKLPLNSQDTEALRASLLAANITA 1079

QY 1079 DLKVPKGTDPAPFVFPYITNVPYQYLTLPEGLFMLSCLIVPTFAVSCLLGLDLRS 1138
 Db 1080 DLKVPKGTDPNPEVFPYITNVPYQYLTLPEGLFMLSCLIVPTFAVSCLLGLDLRS 1139
 QY 1139 LNLISLVMILVDTGFMALWDISYNAVSLINLVSAGVMSVFPVSHITRSPAISTKPTWL 1198
 Db 1140 LNLISLVMILVDTGFMALWDISYNAVSLINLVSAGVMSVFPVSHITRSPAISTKPTWL 1199
 QY 1199 ERAKEATISMSGSAVAGVAMTLPGLVLGLAKAQLIOIFPFRMLNLTLLGLLHGLVFL 1258
 Db 1200 ERAKATVFMGSAVAGVAMTLPGLVLGLAKAQLIOIFPFRMLNLTLLGLLHGLVFL 1259
 QY 1259 PVLSYGVDPNPALEQKRAEAAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGI-KG 1317
 Db 1260 PVLSYGVDPNPALEQKRAEAAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGI-KG 1316
 QY 1318 AGAISNPLNNGRQF 1332
 Db 1317 ANAARSSLPKSDQKF 1331

RESULT 13

AE93568 ID AE93568 standard; protein; 1331 AA.

AC AE93568;

DT 06-OCT-2005. (first entry)

XX Rat NPC1L1 protein.

DE Protein engineering; NPC1L1; antidiabetic; antiarteriosclerotic;
 KW cardiant; vasotropic; hyperlipidemia; atherosclerosis; stroke;
 KW arteriosclerosis; rat.

XX Rattus sp.

XX WO2005069900-A2.

XX 04-AUG-2005.

XX 14-JAN-2005; 2005WO-US001469.

XX 16-JAN-2004; 2004US-0537341P.

XX (MERI) MERCK & CO INC.

XX Garcia-Calvo M;

XX WPI; 2005-564070/57.

XX N-PSDB; AE93567.

XX Identifying ligand of NPC1L1 for stimulating the activity of NPC1L1, by
 PT contacting NPC1L1 with detectably labeled substituted 2-azetidinone
 PT glucuronide and a candidate compound and determining if compound binds to
 PT human NPC1L1.

PS Example 2; SEQ ID NO 2; 215pp; English.

XX The invention relates to identifying a ligand of NPC1L1. The method
 CC involves contacting human NPC1L1 with a detectably labeled substituted 2-
 CC azetidinone glucuronide and a candidate compound and determining whether
 CC the candidate compound binds to human NPC1L1. In identifying a ligand of
 CC NPC1L1, the KD value is 100 nM or lower, preferably 50 nM, 20 nM, or 10
 CC nM or lower. The detectably labeled substituted 2-azetidinone glucuronide
 CC is 35 S-labeled compound 2, given in the specification. NPC1L1 ligands
 CC are useful for stimulating or blocking the activity of NPC1L1, and for
 CC treating conditions caused or mediated by NPC1L1. It is useful for
 CC reducing the incidence of hyperlipidemia, atherosclerosis, coronary heart
 CC disease, stroke, or arteriosclerosis. The present sequence represents a
 CC rat NPC1L1, a N-glycosylated protein.

SQ Sequence 1331 AA;

Query Match 78.5%; Score 5421.5; DB 9; Length 1331;
 Best Local Similarity 77.8%; Pred. No. 0;
 Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
 QY 1 MAEAGRLWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
 Db 1 MAEAGRLWLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 59
 QY 61 TPARKITGDHLLILLOKICPRLTYTGNTQ-ACCSAKQLVLSLEASITKALLTRCPACSDN 119
 Db 60 TPARKITGDHLLILLOKICPRLTYTGNTQ-ACCSAKQLVLSLEASITKALLTRCPACSDN 119
 QY 120 FVNLHCHNTCPNOSLFNVTNRVAQAGAQLPVAVYEAIFYOHSFAEOSYDSCSRVRVPA 179
 Db 120 FVNLHCHNTCPNOSLFNVTNRVAQAGAQLPVAVYEAIFYOHSFAEOSYDSCSRVRVPA 179
 QY 120 FVNLHCHNTCPNOSLFNVTNRVAQAGAQLPVAVYEAIFYOHSFAEOSYDSCSRVRVPA 179
 Db 120 FVNLHCHNTCPNOSLFNVTNRVAQAGAQLPVAVYEAIFYOHSFAEOSYDSCSRVRVPA 179
 QY 180 AATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDTTFLHLEPGQAVSGIQLNBSG 239
 Db 180 AATLAVGTCMGVYGSALCNAQRWLNFGQDTGNGLAPLDTTFLHLEPGQAVSGIQLNBSG 239
 QY 240 VARNESQDDVATCSCODCAASCPAIARPOALDSTFTLQMPGSLVLIILLCSPFVAVT 299
 Db 240 VARNESQDDVATCSCODCAASCPAIARPOALDSTFTLQMPGSLVLIILLCSPFVAVT 299
 QY 300 ILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQPPQGGTVMVSWPLTILVLS 359
 Db 300 ILLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQPPQGGTVMVSWPLTILVLS 359
 QY 360 VIVVALAAGLVFTTETDTPVELMSAPNSQARSEKAFDQHPGPPFRNTQVILTAAPNRSS 419
 Db 360 VIVVALAAGLVFTTETDTPVELMSAPNSQARSEKAFDQHPGPPFRNTQVILTAAPNRSS 419
 QY 420 YRYDSSLILGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSQDICYAPLNPDN 479
 Db 420 YRYDSSLILGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNLSQDICYAPLNPDN 479
 QY 480 TSLYDCCINSILQYFQNNRTLLLTANOTLMGQTSQVDKQHFYLCANAPLTFKDGTALA 539
 Db 480 TSLYDCCINSILQYFQNNRTLLLTANOTLMGQTSQVDKQHFYLCANAPLTFKDGTALA 539
 QY 540 LSCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAALWBEAFLE 599
 Db 540 LSCMADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAALWBEAFLE 599
 QY 600 MRAFORMAGMPQVTFPABRSLEDEINRTTADLPFATSYIIVIFLYTSLALGSYSWSR 659
 Db 600 MRAFORMAGMPQVTFPABRSLEDEINRTTADLPFATSYIIVIFLYTSLALGSYSWSR 659
 QY 660 VMVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVWVPLVLSVGNADNIFPVL 719
 Db 660 VMVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVWVPLVLSVGNADNIFPVL 719
 QY 720 EYQRLPRPGRPREVHIGRALGRVAPSMLLCSLSEACFFFLGALTTPMPAVRTFALTSGIA 779
 Db 720 EYQRLPRPGRPREVHIGRALGRVAPSMLLCSLSEACFFFLGALTTPMPAVRTFALTSGIA 779
 QY 780 VILOFLQMSAPVALLSLDSKRBASRLDVCCCKVKBQLPPPGQEGLLLGFFQKAYAPF 839
 Db 780 VILOFLQMSAPVALLSLDSKRBASRLDVCCCKVKBQLPPPGQEGLLLGFFQKAYAPF 839
 QY 840 LLHWTITGVVLLFLAFGVSLYSNCHTSVGLDDELALPKDSYLLDYLFLNRYPEVGPAP 899
 Db 840 LLHWTITGVVLLFLAFGVSLYSNCHTSVGLDDELALPKDSYLLDYLFLNRYPEVGPAP 899
 QY 900 VYFVTTGYNFSSBAGMNAICSSACGNFSTQKIQTATEPPEQSYLEIPASSWDDFD 959
 Db 900 VYFVTTGYNFSSBAGMNAICSSACGNFSTQKIQTATEPPEQSYLEIPASSWDDFD 959
 QY 960 WLTP-SSCCRLYISGPNKDKFCPSVTNSLNCILKNCMSITMGSVRPSVQFHKYLPWFLND 1018
 Db 960 WLTP-SSCCRLYISGPNKDKFCPSVTNSLNCILKNCMSITMGSVRPSVQFHKYLPWFLND 1019

QY 1019 RPNKCPKGGGLAAXSTSVNLSDGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITA 1078
DB 1020 PPNIRCPKGGGLAAXRTSVNLSSDGOVIASQPMAYHKPLRNSQDPTALRASRLAANITA 1079
QY 1079 DLKVPKGTDAFAEVPYITNVFEOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
DB 1080 DLKVPKGTDPNFEVFPYITNSVFQOYLTVLPEGIFTALCFVPTFVVCYLLGLDMCSG 1139
QY 1139 LNLISIVMILVDTVGFMAWDISNAVSLNLSAVGMSVEFVSHITRSPAIKTKPTWL 1198
DB 1140 ILNLSIIMILVDTLGLNAVGISNAVSLNLSVAVGMSVEFVSHITRSPAVSTKPTRL 1199
QY 1199 ERAKATISMSAGVAGVAMTNLPGLVGLAKAQLIQIFPRLNLLITLGLHLGLVFL 1258
DB 1200 ERAKATVFMGSAVAGVAMTNFPGILITLGAQQLIQIFPRLNLLITLGLHLGLVFL 1259
QY 1259 PVILSVGPDVNPALALQKRAEAAVAVMVASCPNHPHSRVSTADNIVVNSHFEFSI-KG 1317
DB 1260 PVVLSYLGPDVNOALVQBEKLAESA-AVAPBSCPOYPSPADADAN---VNYGFAPELAHG 1316
QY 1318 AGAISNIFLNNRQF 1332
DB 1317 ANAARSSLPKSDQKF 1331

RESULT 14

ID ADJ27305
XX ADJ27305 standard; protein; 1333 AA.

AC ADJ27305;

DT 20-MAY-2004 (first entry)

XX Mouse NPC1L1.

XX Niemann-Pick disease; type C1; gene-like 1; NPC1L1; trans-golgi network;
KW plasma membrane; transport signal; promoter;
KW sterol regulated element binding protein 1; SREBP1;
KW binding consensus sequence; transmembrane domain; sterol-sensing domain;
KW SSD; cholesterol; NPC1; receptor; Niemann-Pick C1 disease; intestinal;
KW cholesterol absorption; serum cholesterol; hyperlipidaemia;
KW atherosclerosis; coronary heart disease; stroke; arteriosclerosis.

XX Mus sp.

XX WO2004009772-A2.

XX 17-JUL-2003; 2003WO-US022467.

XX 19-JUL-2002; 2002US-0397442P.

XX (SCHE) SCHERING CORP.

XX Altman SW, Murgolo NJ, Wang LQ, Graziano MP;

XX WPI; 2004-132945/13.

XX N-PSDB; ADJ27304.

XX New Niemann-Pick disease, type C1, gene-like 1 polypeptide, for detecting
PT or identifying antagonists of NPC1L1 for inhibiting intestinal
PT cholesterol absorption in a subject, or for treating elevated serum
PT cholesterol or stroke.

XX Claim 1; SEQ ID NO 12; 125pp; English.

XX This sequence represents a Niemann-Pick disease, type C1, gene-like 1
CC (NPC1L1) polypeptide. NPC1L1 is an N-glycosylated protein which contains
CC a motif ADJ27331 which acts as a trans-golgi network to plasma membrane
CC transport signal, and which exhibits limited tissue distribution and
CC gastrointestinal abundance. The human NPC1L1 promoter sequence contains a

CC sterol regulated element binding protein 1 (SREBP1) binding consensus
CC sequence. NPC1L1 has 13 transmembrane spanning segments and a sterol-
CC sensing domain (SSD) which is involved in sensing cholesterol levels.
CC possibly by a mechanism which involves direct cholesterol bonding. NPC1L1
CC has 42% amino acid sequence homology to human NPC1, a receptor
CC responsible for Niemann-Pick C1 disease. The NPC1L1 polypeptides or
CC polynucleotides are useful for detecting or identifying antagonists of
CC NPC1L1, which can be used for inhibiting intestinal cholesterol
CC absorption in a subject, or for treating medical conditions including
CC elevated serum cholesterol, hyperlipidaemia, atherosclerosis, coronary
CC heart disease, stroke or arteriosclerosis.
XX
SQ Sequence 1333 AA;

Query Match 78.3%; Score 5407; DB 8; Length 1333;

Best Local Similarity 77.2%; Pred. No. 0;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGLLWALLRLAQSEPVYTIHQPGYCAFYDECNKPGLSMTLSNVSCLSNTPA 63

DB 3 AAWGMLLWALLNLSAQGLYTPHKGACTFYEECKNPELGGTLTSLNISCLNTPA 62

QY 64 RKITGDHLILQKICPLRYTGNP-TQACCSAKQLVLSBASLITKALLTRCPACSDNPNV 122

DB 63 RHVTGDHLALLQRCPLRYNGPNDTYACCSKQLVSLDSSLITKALLTRCPACSENPFV 122

QY 123 LHCHTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFVQHSFAEOSYDSCSRVRVPAAT 182

DB 123 IHCHTCSPPQSLFINVTRVVRQDPQLPVAVVAEAFYQSFABKAVESCRVRIPAAS 182

QY 183 LAVGTMGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLEPQAVGSGIOPNEGVAR 242

DB 183 LAVGSMGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLEPQALADGKMDGKITP 242

QY 243 CNESGDDVATCSQDCQCAASCPAIARPOALDSTVILGOMPGLVLIILCSVFVAVTLL 302

DB 243 CNESGQSDSAACQDCQCAASCPVPPPPALRPSFYMGRMPGMLALIIIFTAVFVLLSVL 302

QY 303 VGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSVP 362

DB 303 VTLRVASNRNKKVAGSQEAPNLPRKRFSPHTVLGRFFESWGTVASWPLTILVLSFIV 362

QY 363 VVALAAGVITFELTTPDVELMSAPNSQARSEKAFHDQHFPGFFRTNQVILFAPNRSRY 422

DB 363 VIALSVGLTFIELTTPDVELMSAPKSQARKEKAFHDEHFGFFRTNQIEVFAKNRSYKY 422

QY 423 DSSLGPKVFSGLDLDLLELLELOERLRHLQVWSPQAQRNISIQDICYAPLNPDNTSL 482

DB 423 DSSLGPKVFSGLDLDLLELLELOERLRHLQVWSEHAQRNISIQDIYAPLNPHNTSL 482

QY 483 YDCCINSLLQVFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTAALASC 542

DB 483 TDCCVNSLLQVFQNNHTLLLTANQTLNGQTSVDMKDFLYCANAPLYKDGTAALASC 542

QY 543 MADYGAPVFPFLAIGGYKGOYSEAEALIMTFSLNYPAGDPRLOAKLWEAEPLSEMPA 602

DB 543 IADYGAPVFPFLAVGGYQGTDYSEAEALITITINNYPADDDPRMAHKLWEAEPLKEMOS 602

QY 603 FORRMAGMFOVTFTAERSLEDEINETTAEGLPIFATSYIVIFLYISLALGYSYSSRVNV 662

DB 603 FORSTADKFOIAFSAERSLEDEINETTQDLPEVFAISYIVIFLYISLALGYSYSSRVNV 662

QY 663 DSKATLGLGGVAVLGVAMMAAGFFSYLGRSSVLQVVPFLVLSVGADNIFIFVLEYQ 722

DB 663 DSKATLGLGGVAVLGVAMMGFFSYLGVPSLSVIQVVPFLVAVGADNIFIFVLEYQ 722

QY 723 RLPRRPGEPREVIHIGRALGRVAPSMMLCSLSAICFFLIGALTMPAVRFTALTGLAVIL 782

DB 723 RLPRMPGEQREAHIGRTLGSAVAPSMMLCSLSAICFFLIGALTMPAVRFTALTGLAIF 782

QY 783 DFLQMSAVFALLSLDLSKQKQASRLDVCCKVPQELPPPGQEGGLLGFQKAYAPFLHL 842

DB 783 DFLQMTAFVALLSLDLSKQKQASRPDVCPCPSRNLPPPKQEGGLLCCFFRKYITPFLHL 842

Db	423	DSLLGPKPNSGTLSDLLQELLEQLERLHQVWSHEAQRNISQDICYAPLNPHTSL	482
Qy	483	YDCINSLLQYFQNNRTHLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTTALASC	542
Db	483	TDCVNSLLQYFQNNHTLLLTANQTLNGQTSLVDMKDHFLYCANAPLTYKDGTTALASC	542
Qy	543	MADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNPNYPAGDRLQAOKLWEBAFLEEMRA	602
Db	543	IADYGAPVFPFLAVGGYQGYDYSEAEALITITSINNPADDPRMMAHAKLWEBAFLKEMOS	602
Qy	603	FORMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYIVIFLYTSLAGSVSSSRVMV	662
Db	603	FORSTADKQIFAFSAERSLEDEINRTTIQDLPFVAISYIVFLYISLAGSVSRMSRVAV	662
Qy	663	DSKATLGLGGAUVVLGAVMAAGFFSYLGRSSLVILQVVPFLVSLVSGADNIFIFVLEYQ	722
Db	663	DSKATLGLGGAUVVLGAVMAAGFYSYLVGPSSVLIIQVVPFLVAVGADNIFIFVLEYQ	722
Qy	723	RLPRRPOGEPREVHIGRALGRVAPSMLLCSLSEAI CFPLGALTPMPAVRTPALTSGLAVTL	782
Db	723	RLPRMPEQREAHIGRTLGVAVPSMLLCSLSEAI CFPLGALTSMPAVRTPALTSGLAII F	782
Qy	783	DELLQWASFAVALLSLDSKROEASRLDVCCCKQDELPPQGOGLLGLFFOKAYAPFLIH	842
Db	783	DFLQMTAFVALLSLDSKROEASRPDVCCFSERNLPPPKQKESLLLCFRKIYTPFLIH	842
Qy	843	WITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPAPVYF	902
Db	843	RFIREPVLLLFVLFGANLYLMCNISVGLDQDLALPKDSYLLIDYFLFLNRYLEVGPVYF	902
Qy	903	VTTLYGNPSSBAGHVAICSSAGCNPSFTOKIQYATEFPFQSYLAIPASSWDDDFDMLT	962
Db	903	DTTSGYNFSTEAGHVAICSSAGCESFSLTKIQYASEFPNQSYVAATASSWDDFIDMLT	962
Qy	963	P-SSCCRLXTSGPNKDXFCPSTVNSLNCNKNCSITMGSRVPSVEQPHKLPWFLNDRPN	1021
Db	963	PSSCCRLIYTRGPHKDXFCPSTDTSFNCLKNCMNRTLGPVRPTTEQPHKLPWFLNDTPN	1022
Qy	1022	IKCPKGGLAAYSTSVNLTSGQVLASRFMAYHKLKNSQDYTBALRAARELANITADLR	1081
Db	1023	IRCPKGGLAAYRTSVNLSDDQIIIASQFMAYHKLPRNSQDFTBALRASRLAANITAE LR	1082
Qy	1082	KVPGTDPAFEVFPYTTNVFVEQVLTTLPGLEFMLSCLVPTPAVSCLLIAGLDLRSGLLN	1141
Db	1083	KVPGTDPNFEVFPYTTISNVFQQTLTVLPGEIFTLALCFPTFVVCVLLGLDIRSGILN	1142
Qy	1142	LLSIVMLIVDTVGFPMALWDISYNAVSLINLVSAGMSVEFVSHITRSFALSTKPTWLER A	1201
Db	1143	LLSIIMILVDTIGLMAVNGISYNAVSLINLVTAVGMSVEFVSHITRSFAVSTKPTRLER A	1202
Qy	1202	KEATISGMSAPFAGVAMTNLPGLVLVLGAQAQLTQIEFFRLNLLITLLGLHGLVFLPVI	1261
Db	1203	KDATIFMGSAPFAGVAMTNFFGILITLFGAQAQLTQIEFFRLNLLITLLGLHGLVFLPVV	1262
Qy	1262	LSYVGPDVNPALAEOKRAEAAVAVMVVASCPNHPSRVSTADNITYNVHNSPEGS- IKGAGA	1320
Db	1263	LSYLGPDVNPQALVLEEKATEA-AMVSEPSQCQYFPFPADANTISDYVNYGNFPRIPEINA	1321
Qy	1321	ISNFLPNNGROF	1332
Db	1322	ASSSLPKSDQKF	1333

RESULT 16
AEB93578
ID AEB93578 standard; protein; 1333 AA.

XX	DT	06-OCT-2005	(first entry)

XX
DE Mouse NPC1L1 protein.

Qy	181	ATLAVGTCMGVYGALCNQAWLNFQDGTGNGLAPLDITFHLLPEQAVSGIQLPNEG	240		
Db	181	ATLAVGTCMGVYGALCNQAWLNFQDGTGNGLAPLDITFHLLPEQAVSGIQLPNEG	240		
Qy	241	ARCNESQDDVATCSQDCASCP--AIARQALDSTFYLGOMPGSLVLIILCSFPAV	298		
Db	241	ARCNESQDDVATCSQDCASCPKPAPRPSTPSTW--ROMPGSLVLIILCSFPAV	298		
Qy	299	TILLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVL	358		
Db	299	TILLVGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVL	358		
Qy	359	SVIPVVALAAGLVFTLITDTPVELWAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS	418		
Db	359	SVIPVVALAAGLVFTLITDTPVELWAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS	418		
Qy	419	SYRYDSLGLPKNFGIILDLLELLELQERLRLHQLWSPQARNISLQDICYAPLNPD	478		
Db	419	SYRYDSLGLPKNFGIILDLLELLELQERLRLHQLWSPQARNISLQDICYAPLNPD	478		
Qy	479	NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTAKGTAL	538		
Db	479	NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTAKGTAL	538		
Qy	539	ALSCMADYGAPVPPFLAIGYKGYSEAEALIMTFSLNYPAGDPRLAQAQKWEAEFL	598		
Db	539	ALSCMADYGAPVPPFLAIGYKGYSEAEALIMTFSLNYPAGDPRLAQAQKWEAEFL	598		
Qy	599	EMRAFORRMAGMFOVTFEAERLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSWS	658		
Db	599	EMRAFORRMAGMFOVTFEAERLEDEINRTTAEDLPFATSVIVIFLYISLALGSYSWS	658		
Qy	659	RVWVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSVLIVQVPLVLSVGADNIFIV	718		
Db	659	RVWVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSVLIVQVPLVLSVGADNIFIV	718		
Qy	719	LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEATCFGLALTPMPAVRTALTSL	778		
Db	719	LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEATCFGLALTPMPAVRTALTSL	778		
Qy	779	AVILDFLLQMSAFVALLSDSKQEAASRLDVCCKVQBELPPPGQGEGLLGLFFQKAYAP	838		
Db	779	AVILDFLLQMSAFVALLSDSKQEAASRLDVCCKVQBELPPPGQGEGLLGLFFQKAYAP	838		
Qy	839	FLLHWITRGVW----LLLFLALFGVSLXSNCHISVGLDQELALPK	879		
Db	839	FLLHWITRGVWVPSQLLFLALFGVSLXSNCHISVGLDQELALPK	883		
RESULT 18					
ID	AAW88445	AAW88445 standard; protein; 1278 AA.			
XX	XX				
AC	AAW88445;				
XX	XX				
DT	26-APR-1999	(first entry)			
XX	XX	Human NPC1 (Niemann-Pick type C disease) polypeptide.			
DE	XX				
XX	XX	Niemann-Pick disease type C; NPC1 gene; human; diagnosis; therapy;			
KW	KW	cholesterol; neurodegeneration.			
XX	OS	Homo sapiens.			
PH	' Key *	Location/Qualifiers			
FT	Peptide	1. .24			
FT		/note= "signal peptide"			
FT	Protein	25. .1278			
FT		/note= "signal peptide"			
FT	Region	55. .164			
FT		/note= "conserved region"			
FT	Region	73. .94			
FT		/note= "leucine heptad motif, or leucine zipper"			
FT	Peptide	1275. .1278			
FT		/note= "lysosome taretting motif"			
PN	WO9901555-A1.				
XX	XX				
PD	14-JAN-1999.				
XX	XX				
PF	02-JUL-1998;	98WO-US013862.			
XX	XX				
PR	03-JUL-1997;	97US-0051682P.			
XX	XX				
PA	(USSH)	US DEPT HEALTH & HUMAN RESOURCES.			
XX	XX				
PI	Carstea ED, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;				
PI	Rosenfeld MA, Loftus SK, Gu J;				
XX	XX				
DR	WPI; 1999-106056/09.				
DR	N-PSDB; AAX06873.				
XX	XX				
PT	New isolated gene, NPC-1 - is associated with Niemann-Pick type C				
PT	disease, used to develop products for the study, diagnosis and therapy of				
PT	the disease.				
PS	Claim 1; Page 56-60; 101pp; English.				
XX	XX				
CC	This polypeptide is encoded by the human NPC-1 gene (see AAX06873) that				
CC	is associated with Niemann-Pick disease type C (NP-C). NPC1 when				
CC	expressed in human cells derived from NP-C patients is able to correct				
CC	the abnormal lysosomal cholesterol accumulation that is characteristic of				
CC	such cells. Hydrophobicity studies and structural motif comparisons				
CC	suggest the polypeptide is an integral membrane protein with as many as				
CC	16 possible transmembrane regions. Database sequence comparisons reveal				
CC	extensive identity/similarity to NP-C orthologues in mouse, yeast and				
CC	Caenorhabditis elegans (see AAW88446-48). The provision of the human NPC1				
CC	cDNA sequence enables methods of detecting the presence of mutations in				
CC	the gene corresponding to the hNPC1 cDNA, and thereby facilitates the				
CC	determination of whether an individual is an NP-C sufferer, or an NP-C				
CC	carrier. The products can also be used for the study of NP-C and in				
CC	therapeutic applications related to NP-C				
XX	XX				
SQ	Sequence 1278 AA;				
Query Match		34.8%;	Score 2402.5; DB 2; Length 1278;		
Best Local Similarity		39.8%;	Pred. No. 1.8e-214;		
Matches 522;		Conservative 237;	Mismatches 450; Indels 103; Gaps 25;		
Qy	7	RGLWLLALLLRLAQSEPYTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVCLS	59		
Db	4	RGLALGLLLLLLCPAQVFSQ-----SCVMYGECCGIAYGDKRYNCEYSG-----	46		
Qy	60	NTPARKITGDHLLILQIKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDN	119		
Db	47	--PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKNTLQPLQFLSRCPSCFYN	103		
Qy	120	FVNLHCHNTCSNQSFLINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAESQSDSCSRV	175		
Db	104	LLNLFCELTCSPRQSQFLNVTATEDYVDPVTNTQTNVKELQYVYGQSPANAMYNACRDV	163		
Qy	176	RVPAATATLAVGTCMGVYGALCNQAWLNFQDGTGNGLAPLDIT-----FHLLPEGOA	228		
Db	164	EAPSSNDKALGLLCKGRADA-CNATNWIYFMFNKNGQAPFTITPVFSDFPVH-----	215		
Qy	229	VSGGIQPLNEGVARCNESQDDVATCSQDCASCPAIARPO-----ALDSTFYLG	279		
Db	216	---GMEPMNATKGCDESVDVTAPCSQCDCSIVCGPKPQPPPPAPWPTILGLDAMYIM	272		
Qy	280	QMPGSLVLIILCSVPAVVITLL-----VGFVAPARDKSKMVDPKKGTSL	326		
Db	273	WITYMAFLLVFFGAPFAVWCKRYKRVFSYTYTIDSNIAFSV-NASDKG-----ASCC	324		
Qy	327	DKLSFSTHTLLGQFFQGWGTWASWPLTILVLSLIPVVALAAGLVFTLTDVPELWAP	386		
Db	325	DPVSAAFEGCLRLFRNGSCFVRNPGCVIFFSLVFITACSSGLVFRVVTNPDVLSAP	384		

CC associated gene containing one or more SNP's of the invention. Note: This sequence was not shown in the specification. The sequence has come from an electronic sequence listing downloaded from the WIPO website.

XX * Sequence 1278 AA;

Query Match 34.8%; Score 2402.5; DB 8; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.8e-214;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWALLRLAQSPPYTHIQPGYCAFVDECG-----KNPELSSGSLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVYGEIGIAYGKRYNCEYSG-----46
QY 60 NTPARKITGDHLILLOKICPRILYTGENTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGLVQELCFGFFFG-NVSLCCDVRQLQTKDLQPLQFLSRCPFCFN 103
QY 120 FVNLCNTCSNQSFLINWTR-----VAQLGAGQLPAVAYEAFYQHSFABOSYDSCSRV 175
DB 104 LNLFLCELTCSPRSQFLNVTATEDYVDFVTNQTKTNKVELQYVVGQSPANAMYNACRDV 163
QY 176 RYPAATLAVGTMGVYGSAICNAQRWLNFGQDGTGNGLAPLDT-----PHLEPGQA 228
DB 164 EAPSNDRKALGGLCGKADA-CNATNWIEMFNKONGQAPFTITPVFSDFPVH-----215
QY 229 VSGSQPLMEGVARNCSOGDDVATCSQDCACSPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGDSDVETAPCSQDCSIVCGPKPQPPPPAPWTLGLDAMYVM 272
QY 280 QMGSLVLIILCSVPVAVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
DB 273 WITYMAFLLVFGAFVAVCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSSTHTLLQFQCGWGTWASPLTLVLVSVPVVALAAGLVFETLTPDVELWSAP 386
DB 325 DVPSAAFEGLRLFRWGSFCVRNPGCVIFPSLVFETACSGLVFVRVTTNPFVLSAP 384
QY 387 NSQARSEKAFHQHGFPPRTNQVILTAPNRSYRYSLLGPK-NFSGILDLDLLELL 445
DB 385 SQAARLEKEYFOHGFPPFTQLIIRAPLTKHIYQPYPSGADVPFGPPLDIQILHQVL 444
QY 446 ELQERLRLHQQVSPARQNISLQICYAPLNPNTSLYDCCNSLQYQNNRTLLLLTA 505
DB 445 DIQIAIEN--ITASVDNETVTLQDLCAPLSPYNT--NCTILSVLNYFQNSHSLDHKK 499
QY 506 NQTLMGQTSQVWDKHFLYCANAPLTFKDGTLALSCMADYGAPEPFLAIGYKGDYS 565
DB 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPCLGTGCGPFPPLVLGQYDDQNYN 556
QY 566 EBAELIMTFLSNYPAGDPLRLAOKLWBEAFLEEMRAFORRMAGMFQVTTFAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKQRAQWKEPINFVKNYKN---PNLTISFTAERSLEDEI 613
QY 626 NRTTAEDLPATYSIVIFLYISLALGYSYSSRVMSVDSKATILGGLGVAVVLGAYMAAG 685
DB 614 NRESDSVFTVISYAIMFLYISLALGHKSCRELLVDSKSLGIAGIILVLSVACSIG 673
QY 686 FFSYLGIRSLVILQVPPFLVSVGADNIPFVLEYQRLPRRPGPREVHIQALGRVAP 745
DB 674 VFSYGLPLTLVIEVLPVLAVGVNDNIFILVQAYQRDERLQGETLQDLQGLRVLGEVAP 733
QY 746 SMLCGLSLEAICFFLGALTPMAVTEALTSLAVLDELQMSAPVALLSDSKROAS 805
DB 734 SMLFSFSFTVAFGLSVMFAVHTFSLFAGLAVFIDPLQLITCFVSLGLGDIKRQEN 793
QY 806 RLDCVCCVQBLPPPGQ-GEGLLGFQKAYAPFLHWTIRGVLALLFALFGVSLYSM 864
DB 794 RLDFCCVGRGAEDGTSVQASECLPFFFKNSYSPILLKDMWPIVIAIFVGLVSLIAVL 853
QY 865 CHISVGLDQELAPKDSYLLDYLFLNRYFVEGAPVYFVTLLGYNFPSSBAGNAICSSAG 924
DB 854 NKVDIGLQSLMPDSDVMVDYFKSISQYLHAGPPVYFVLEBHDYTSKQGMVCGMG 913

QY 925 CNFSTFTQKIQYATBPPEQSYLAIPASSWVDQFIDWLTTP-SSCCRLYISGPNKDKFCPST 983
DB 914 CNNDLSVQOIFNAQAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCNKCMSIT-MGSVRPSVBQPHKYLPHFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VYDPACVR-CRPLTPEGKQPGQDFMRFLPMLFSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQCVLASRPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTNV 1100
DB 1030 GTRVGATYPTWTHVLTQTSADFIDALKARLLASNT-ETMGINGS--AYRFPYSPVYV 1086
QY 1101 FTEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSYVMILVDTVGFNALMD 1160
DB 1087 FTEQYLTIIIDDTIFNLGVSIGAIPLVMVLLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146
QY 1161 ISYNAVSLINLYSAVMSVEFVSHITRSPAISTKTWLERAKETISMGSAPVAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERABEALAHMGSSVFSGITLTK 1206
QY 1221 LPGILVGLAKAQLIQIPFRNLNLTITLGLLGLVFLFVLVLSYVCGPDVNP 1272
DB 1207 FGGIVVLPAPAKSQIFQIFFRMLAMVLGATHGLIFLPLVLLSYIGPSVYKA 1258

RESULT 21

ADU06723

ID ADU06723 standard; protein; 1278 AA.

XX AC ADU06723;

XX XX 27-JAN-2005 (first entry)

XX DE Novel bronchial cancer-associated human protein SeqID949.

XX KW bronchial cancer; cytostatic; tumour-associated protein;

XX KW cancer detection; metastasis; tumour; human.

XX OS Homo sapiens.

XX XX DE10316701-A1.

XX PD 04-NOV-2004.

XX PF 09-APR-2003; 2003DB-01016701.

XX PR 09-APR-2003; 2003DB-01016701.

XX PA (HINZ/) HINZMANN B.

XX PA (HERM/) HERMANN K.

XX PA (CAST/) HEIDEN CASTANOS-VELEZ E.

XX PI Mennerich D, Bruemendorf T, Heiden E, Hermann K, Kinnemann H;

XX PI Li X, Roepcke S, Staub E, Hinzmann B, Rosenthal A, Pilarczyk C;

XX DR WPI: 2004-786403/78.

XX DR N-PSDB; ADU06236.

XX PT New nucleic acid, and derived proteins, useful for diagnosis of bronchial cancer and in screening for therapeutic and diagnostic agents.

XX PS Claim 2; SEQ ID NO 949; 1381pp; German.

XX CC This invention relates to a novel isolated nucleic acid associated with bronchial cancer comprising 489 defined sequences given in the specification. The invention may be useful for the production of compounds with a cytostatic activity through the inhibition of expression or activity of tumour-associated proteins. The novel DNA sequences and the proteins/peptides encoded by them are used for detecting bronchial cancer or determining the risk of developing it and to screen for specific binding partners of the DNA or protein sequences, where the binding partners are potentially useful as agents for treating or

CC diagnosing bronchial cancer. The DNA or protein sequences can also be
CC used for prognosis, detection of metastases and for secondary treatment
CC (of tumours that have been stabilised or are no longer detectable).
CC Detecting abnormal expression of the DNA sequences provides early
CC diagnosis of bronchial cancers. The present sequence is that of a protein
CC encoded by a novel bronchial cancer-associated human gene sequence of the
CC invention.

XX SQ Sequence 1278 AA;
Query Match 34.8%; Score 2402.5; DB 8; Length 1278;
Best Local Similarity 39.8%; Pred No. 1.8e-214;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 RGLWALLLRLLAQSEPTTHIQGYCAPYDECG-----KNPELSGLMTLSNVSL 59
DB 4 RGLALLGLLLCPAQFSQ-----SCVMYGCGLAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLLQKICPRLYTGPTQACCSAKOLVLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLQELCPGFFG--NVSLCCDVRQLQTLKDNLQLPLQFUSRCPCFYN 103
QY 120 FYNLCHNTCSNQSFLINVT---VAQLGAGQLPAVVAYEAFYQHSFAEQSYDCSRV 175
DB 104 LLNLFCELTCSRQSGFLNVTATEDYDPVTQTKTNKELQYYVQSFANMYNACRDV 163
QY 176 RVPAAATLAVGTCGYGSGALCNAQRWLNFOQDGTNGNLAPLDIT-----PHLEPQA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPPTITPVFSDFVH----- 215
QY 229 VSGIOPLEGVARNCHESQGDVATCSQDCAASCPAIAAPQ-----ALDSTFYLG 279
DB 216 --GMEPMNATKGCDSDEVTAPCQCDCSIVCGKFPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPSGLVLIILCSFVAVTILL-----VGRVAPARDKSMVDPKKGTSL 326
DB 273 WIYMAFLVFFGAFVWCYKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGGWTWASWPLTILVLISVIVVLAAGLVTELTTPDVLWSAP 386
DB 325 DPVSAAPFEGCLRLFRWGSFCVRNPGCVIFFSLVFITACSSGLVFPVRVTTPDVLWSAP 384
QY 387 NSQARSEKAFHQHGFPEFTNVOILTAPNRSRYRDSLGLPK-NFSGILLDLLELL 445
DB 385 SSQARLEKXYDQHGFFFEQLIRAPLTDKHIYQPYPSGADVPFGPLDIQLHQVL 444
QY 446 ELQERLRLQVMSPEAQRNLSIQDICVAPLNPONTSLYDCCINSLQYFQNNRTLLLP 505
DB 445 DLQIAIEN--ITASYNETVTLQDLCIAPLSPYNT--NCTILSVLNYFQNSHSLDHKK 499
QY 506 NQTLMGOTSQVNDKDHPLCANAPLTFKQGTALALSCMADYCAPVPFFLAIGYKGDYS 565
DB 500 GDDPFF--VYADVHTHELYCVRAPASLNDTSLHDFCLGTGFGVPFVFWLVLGGYDDQYN 556
QY 566 EAEALIMTFSLNNYPAGDPRLAQAKWEEAFLEEMRAFORRMAGMVFQVTFATERSLEDEI 625
DB 557 NATALVTFPVNNYNDTEKLQRAQAWKEKFINFVKNKN---PNLTISTAERSIEDEL 613
QY 626 NRTTAEDLPFATSYIYIFLYISALGSYSWSRVMVDSKATLGLGAVVVLGAVMAAMG 685
DB 614 NRESDSVFTVWISYAIMFLYISALGHKICRLLVDSKVSLGIAGILIVSSVACSLG 673
QY 686 FFSYLGTRSSLVLQVVPFLVSGADNIIFIULYQRLPRRPGEPREVIHGRALGRVAP 745
DB 674 VFSYIGLPLTLIVIEVFPFLVAVGVNDNIIFILVQAYQORDERLQGETLDQQLGRVLGEVAP 733
QY 746 SWLLCSLSEATCFPLGALTWPAVRTFALTGSLAVILDLLQMSAFVALLSLDSKQREAS 805
DB 734 SMFLSSFSSETVAFPLGLASWPAVHTFSLFAGLAVFDILQITCFVSLGLGLDKQKKN 793
QY 806 RLDVCCCKVKPQLPFPQ--GEGLLGFFQKAYAPFLHMTIRGVVLLFLALFGVSLYSM 864
DB 794 RLDIFCCVRGAEDGTSVQASBSCLFRFFKNSYSPLLKDWMRPIVIAIFVGLVLSFAVL 853

QY 865 CHLSVGLDQELALPKDSYLLDYFLNRYPEVGCAPVYVFTTLGYNFSSAGMNAICSSAG 924
DB 854 NKVDIGLDQSLSPDDSYWVDYFKSISQYLAHAGPPYFVLEEGHDYTSKQGNVCCGMG 913
QY 925 CNNFSFTQKIQTAYTEPPEQSYLAIPASSWVDDFIDMLTP--SSCRLYISGPNKDKPCPST 983
DB 914 CNNDLSVQQIFNAQAQLDNYTRIGFAPSSWIDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNCILKNCMSIT--MGSVPSVEQPHKYLPMELNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VVDPACVR--CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGRGAHAYSSAVNILLCH 1029
QY 1041 DGQVLASREMAXKPKLKNQSDYTEALRAARELANITADLRKVPGTDPAPAEVFPYTTIN 1100
DB 1030 GTRVGATYFTYHTVLTQTSADFIDALKARLIASNVT--ETMGINGS--AYRVFPYSVFYV 1086
QY 1101 FYEQYLTILPEGLPMLSLCLVPTFAVSVCLLLGLDLRLSGLNLNLSIVMLVDVTGFMALWD 1160
DB 1087 FYEQYLTIIODTIFNLGVSGLAIFLVTMVLGCELWSAVIMCATIAMVLMVNMFGVMLWG 1146
QY 1161 ISYNAVSLINLVSAGVSHITRSPALSTRKPTWLERAKEATISMGSAPVAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSCGISVBFCSHITRAFTVSMKGSVRERABEALAHMGSSVFSGIIITLK 1206
QY 1221 LPGAIVLGLAKAQLIQIFFFRNLILLITLLGLLHGLVFLPVLISYVGPDPNPA 1272
DB 1207 FGGIVLAFAKSQIFQYFMYLAWLLGATHGLIFLPLVLSYIGPSVKA 1258
RESULT 22
AAW88446
ID AAW88446 standard; protein; 1319 AA.
XX AC AAW88446;
XX DT 26-APR-1999 (first entry)
XX DE Mouse NPC1 orthologue.
XX XX Niemann-Pick disease type C; NPC1 gene; mouse; orthologue; diagnosis;
KW therapy; cholesterol; neurodegeneration.
XX OS Mus sp.
XX PN WO9901555-A1.
XX PD 14-JAN-1999.
XX PF 02-JUL-1998; 98WO-US013862.
XX PR 03-JUL-1997; 97US-0051682P.
XX PA (USSH) US DEPT HEALTH & HUMAN RESOURCES.
XX PI Carstee ED, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;
PI Rosenfeld MA, Loftus SK, Gu J;
XX WPI; 1999-106056/09.
XX DR N-PSDB; AAX06874.
XX PT New isolated gene, NPC-1 - is associated with Niemann-Pick type C
XX PT disease, used to develop products for the study, diagnosis and therapy of
XX PT the disease.
XX PS Claim 1; Page 66-70; 101pp; English.
XX CC This polypeptide comprises the murine orthologue of human NPC-1 (see
CC AAW99445) that is associated with Niemann-Pick disease type C (NP-C). The
CC amino acid sequence, predicted from an isolated cDNA clone (see
CC AAX06874), comprises a putative N-terminal signal peptide followed by a
CC domain that is unique to the NPC1 orthologues (in mouse, human,
CC Caenorhabditis elegans and Saccharomyces cerevisiae, see AAW88445-48),

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from *Drosophila*. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (ABBS7737-ABBS72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published/pct sequences

Qy	860	SLYSMCHISVGLDQELALPKDSVLLDYFLFLNRYPEVGAPVYFVTTLLGYNFSEACGMNAI	919
Db	798	SLMVTPISEKGLDQEMSPKNSHVVFYFMYVDLLANGAPVYVWLKPGUNISEPQOQLI	857
Qy	920	CSSAGCNPFSTQKIQYATEFPEQSYLAIPASSWDDFDIWLTPSSCCRLYISGPNKDXF	979
Db	858	CGGVECNNSLSVQLYTOAQYPEITSLARPASSWLLDDYIDLWLAISDCCKYNTV---TGGF	914
Qy	980	CPSTVNSLNLKNCMSITWGSVRPSVEQFHKYLPHFLNDRPNIKCPKGGLAAYSTSVNLT	1039
Db	915	CSSNSKSEDCPCPERGFTENGRLPDAETFNKIPYFLFDLPDAECAKAGRASVADAVIYT	974
Qy	1040	SD-----GQVLASFMYAHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAPFVFPY	1095
Db	975	IDDVGHSTVQDSYPMQYSTTSTSEFYSQLRVREIRISGEINAMPKE---NNVDAEIFAY	1031
Qy	1096	TITNVFEQYLITLPGELFMLSICLVPTFAVSCLLIGLDLRSGLMLLSLVMLLVDTVGF	1155
Db	1032	CVFYIYEQYLITWGDAMFSLGMSLVAFILVTLITGLDITSTFIVLFWVICILINMLGM	1091
Qy	1156	MALWDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAG	1215
Db	1092	MWAWSINLNAISLVNLVVCVGIGVEFVAHLVRSFK-RAEGTAQERARHSLNVTGSSVLSG	1150
Qy	1216	VAMTNLPGLIVLGLAKAQLTIQIFFRINLLITLLGLLHGLVFLPVILSYVGPDPNPALAL	1275
Db	1151	ITLTKPAGIVVLGFSNSQIFQVYFMYGLVIGLGAHGLILLPVLLSLILGPPQK----L	1206
Qy	1276	EQKRAEEAAVAMVVAS	1291
Db	1207	ARSSGAEPTASITIT	1222
RESULT 26			
AAW88447	AAW88447 standard; protein; 1170 AA.		
XX	AC	AAW88447;	
XX	DT	26-APR-1999 (first entry)	
XX	DE	Yeast NPC1 protein orthologue.	
XX	KW	Niemann-Pick disease type C; NPC1 gene; Yeast; orthologue; diagnosis;	
XX	KW	therapy; animal model; cholesterol; neurodegeneration.	
XX	OS	Saccharomyces cerevisiae.	
XX	PN	WO9901555-A1.	
XX	PD	14-JAN-1999.	
XX	PF	02-JUL-1998; 98WO-US013862.	
XX	PR	03-JUL-1997; 97US-0051682P.	
XX	PA	(USSH) US DEPT HEALTH & HUMAN RESOURCES.	
XX	PI	Carstee ED, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;	
XX	PI	Rosenfeld MA, Loftus SK, Gu J;	
XX	XX	WPI; 1999-106056/09.	
XX	DR	N-PSDB; AAX06875.	
XX	PT	New isolated gene, NPC-1 - is associated with Niemann-Pick type C	
XX	PT	disease, used to develop products for the study, diagnosis and therapy of	
XX	XX	the disease.	
XX	PS	Disclosure; Page 75-79; 101pp; English.	
XX	CC	This polypeptide comprises the Saccharomyces cerevisiae orthologue of the	
XX	CC	human NPC1 polypeptide (see AAW88445) that is associated with Niemann-	
XX	CC	Pick disease type 2 (NP-C). The polypeptide shows extensive identity	

CC (34%) and similarity (57%) to the human NPC1 protein. Biochemical and genetic analysis of yeast, worm and murine NPC1 model systems will provide resources for understanding the role of NPC1 in intracellular cholesterol homeostasis and in the aetiology of neurodegeneration in NPC-C disease. The provision of a human NPC1 cDNA sequence (see AAX06873) enables methods of detecting the presence of mutations in the hNPC1 gene, and thereby facilitates the determination of whether an individual is an NPC-C sufferer or carrier

XX Sequence 1170 AA;
SQ

Query Match	19.2%	Score 1329	DB 2	Length 1170
Best Local Similarity	28.8%	Pred. No. 5.7e-114		
Matches 373	Conservative 230	Mismatches 512	Indels 182	Gaps 46
QY	10	LLMAL-----LLRLAQSEPTTHIQPGVCAPVDSCKNP-----ELSGSLMTLSNVCSLNS	60	
DB	3	VLMTIALVQGLUMLVQG-----TATCAMYGNCGKKSVEFNSLPCVPRSPFPPVLSD	54	
QY	61	TPARKTGDHLILKQIKCPRLYTGPNTOQACSAKQLVSLSEASLSITKALLTRCPACSDNF	120	
DB	55	ETSK-----LLVEVGGEEMKEVR-YACCTKDOVVALRDNLQKAPLTISSCPACLKNP	105	
QY	121	VNLCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAPYQHSFABOSYDCSRVRPAA	180	
DB	106	NNLPCHFTCAADQGRFVNITKV-EKSKEDKDIIVAEILDVPMNSSWASEFYDSCNKIKFSAT	164	
QY	181	ATLAVGTMCVGYGSALCNAQRLWLNFGQDTGN--GLAPLDTIFHLLPEQAVGSGIOLPNE	238	
DB	165	N----GYAMDILGGGAKNYSQFLKLGADKPMGLGGSPFQINKYKDLANEE--KEWQBFND	218	
QY	239	GVARCNSQGDVATCSODCAASCAPAIARPOALDSTFYLGQMPG---SLVLIILICSVF	295	
DB	219	EVTACDDAQ---YKCAQSCQESCPHL-KP-LKDGCKVGKPLPCFSLSLVIFTIICALP	272	
QY	296	AVVTILLVGRVAPARDKSMVDPKKGTISLSDK-----LSFSTHT-----LLGQFFQCGW	345	
DB	273	AFWMYYLCKRKGAMIVDDDIIVESG-SLDSESTNVPESENNETNPFNGKLANLFTKVG	331	
QY	346	TWVASWPLTILVSVIPVVALAAGLV-FTELTTPDELWSAPNSQAKSEKAFHDHQGPFP	404	
DB	332	QFSVENPYKILITTVFSIFVFSFIIOVATLETDPINLWVSKNSEKPEKEYPDPDNFGPF	391	
QY	405	FRINOVILTAPNRSSRYDSSL--LLGPKNYSGLDLDLLELLELQELRLHLQWSEPAQ	462	
DB	392	YRTEQIFPVNTEGVLVSYTELHWFDVENE-----ITBEL-----QSS	429	
QY	463	RNTSLQDICIAPLNPDPNTSLDYDCINSLLOYFQNNRITLTLTANQTLMGQTSQVDWKDHP	522	
DB	430	ENTGYQDLCFRP-TEDST----CVIESFTQYFQG-----ALPNKOSWKREL	470	
QY	523	LYCANAPLTFKQGTALALSCMADYGAVPFPPLTAIGYKGYKDYSEABALIMTFSLNNY	582	
DB	471	QEGCKFP-----VNCILPTQEPKLTNLL--FSDDIILNAHAFVVTLLTNNH---	514	
QY	583	DPRLAQAKLWEEAPLEENRAPQORMAGFQVTFPAERLSLEDEINRTTAEDLPATPSYIV	642	
DB	515	---TQSANRWEER--LEBYLLDLKVPGL-:RISFNTEISLEKELNN--NDISTVAISYLM	567	
QY	643	IFLYISIALGSSWSRVMVDSKATILGCGVAVVLGAVMAAMGPFYSVIGRSSIAVILQVW	702	
DB	568	MFLYATWALRRKQKTKLL-----LGISGLIIVLASIVCAAGPLTIFLGRKSTUIIAEVI	621	
QY	703	PFLVLSVGADNIFIIVLEYOR-LPRRPGEPREVHIGRALGRVAPSMLLCSLSSEAICPFLG	761	
DB	622	PFLILAIGIDNIFLI THEYDNCEQKEYSIDOKIISAI GMSFSLMSLLCQTGCFLLA	681	
QY	762	ALTMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKRSQEAASRLDVCCKVPQELPPP	821	
DB	682	AFVTMPAVHNFAIYSTVSIVFNGVLQTLTAVVLSLSEKRSNYKQIT-----	728	
QY	822	QGEGLLGFQKAYAPFLHWTGRGVULLFLALFGVSLYSMCHISVGLDQELALPKDS	881	

729	-GNSETKESFLKTFYFKMLTQ---KRLIIIIISANFPTSIVLPETQIGLDQTLAVPQDS	784
882	YLLDYFLFLNRYFEVGPVFTVTTGLYNFSEAGMNAICSS-AGCNNFSPFKIOYATEF	940
785	YLVDFKDVVYFPLNGPVPVWVK-NLDLFRQNOQKICGKFTTCERDSLANVLB---	840
941	PEOSYLAIPASSWDDPIDMLTPSS--CCRLYISGPNKDKPCPSTVNSLNCNKCMSITM	998
841	RHRSTITEPLANMLDDYFMLNPNQDCCRL---KKGTDEVCVPPSPRRRC-ETCFQ--Q	894
999	GSVR-----PSVEQEHKLPWFLNDRPNIKCPKGLAAYSTSVNLTSDCQVLASRMAY	1052
895	GSWNYNMSPGEGKDFMEFYSIWIN-APSDCPGLGRAPYSTAL-VYNETSVSASVFRTA	952
1053	HKPLKNSQDYTEALRAARELAANITADLRKVPQGTDAPEFVPPYITITNVFYEQVLTILPEG	1112
953	HHPLRSQKDIQAY-----SDGVRISSSPFELDMFAYSPPFYFFVQYQTIGPLT	1001
1113	LFMLSLCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWDISYNVAVSLNV	1172
1002	LKLTGSAIILIFITSSVFL-QNIRSSPFLALVVTMIIVDIGALMALIGISLNAVSLVNL	1060
1173	SAVCMSVEFVSHITRSPAI---STKPTWLERAKEATISMGSAVPAGVAMTNLPGLVIGL	1229
1061	ICVGLGEFCVHIVRSTVVPSETKKDANSRVLYSLNTIGESVIRKGITLTKFIGVCVLAF	1120
1230	AKAQLIQIFPFRNLNLTLLGLLHGLVFLFVILSVYG	1266
1121	AQSKIPDVFFYRWVPTLIIVAAHLALFLPALLSLFG	1157
RESULT 27		
AAW88448		
ID	AAW88448 standard; protein; 1296 AA.	
AC	AAW88448;	
DT	26-APR-1999 (first entry)	
DE	Caenorhabditis elegans NPC1 protein orthologue.	
KW	Niemann-Pick disease type C; NPC1 gene; worm; orthologue; diagnosis;	
KW	therapy; animal model; cholesterol; neurodegeneration.	
OS	Caenorhabditis elegans.	
PN	WO9901555-A1.	
PD	14-JAN-1999.	
PF	02-JUL-1998; 98WO-US013862.	
PR	03-JUL-1997; 97US-0051682P.	
PA	(USSH) US DEPT HEALTH & HUMAN RESOURCES.	
PI	Carstee DA, Tagle DA, Morris JA, Pentchev PG, Pavan WJ;	
PI	Rosenfeld MA, Loftus SK, Gu J;	
XX	WPI; 1999-106056/09.	
DR	N-PSDB; AAX06876, AAX06877.	
XX		
PT	New isolated gene, NPC-1 - is associated with Niemann-Pick type C	
PT	disease, used to develop products for the study, diagnosis and therapy of	
PT	the disease.	
XX	Disclosure; Page 91-95; 101pp; English.	
PS		
XX		
CC	This polypeptide comprises the Caenorhabditis elegans orthologue of the	
CC	human NPC1 polypeptide that is associated with Niemann-Pick disease type	
CC	2 (NP-C). The polypeptide shows extensive identity (30%) and similarity	
CC	(55%) to the human NPC1 protein (see AAW88445). Biochemical and genetic	
CC	analysis of yeast, worm and murine NPC1 model systems will provide	

PT New isolated gene, NPC-1 - is associated with Niemann-Pick type C
PT disease, used to develop products for the study, diagnosis and therapy of
PT the disease.

XX PS Disclosure: Page 91-95; 101pp; English.

This polypeptide comprises the *Caenorhabditis elegans* orthologue of the human NPC1 polypeptide that is associated with Niemann-Pick disease type 2 (NP-C). The polypeptide shows extensive identity (30%) and similarity (55%) to the human NPC1 protein (see AAW89445). Biochemical and genetic analysis of yeast, worm and murine NPC1 model systems will provide

CC resources for understanding the role of NPC1 in intracellular cholesterol
CC homeostasis and in the aetiology of neurodegeneration in NP-C disease.
CC The provision of the human NPC1 cDNA sequence (see AAX06873) enables
CC methods of detecting the presence of mutations in the hNPC1 gene, and
CC thereby facilitates the determination of whether an individual is an NP-C
CC sufferer or carrier
XX
SQ Sequence 1296 AA;
Query Match 15.6%; Score 1078; DB 2; Length 1296;
Best Local Similarity 24.0%; Pred. No. 2.2e-90;
Matches 321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;
QY 76 KTCPLRYTCGNTQACSAKQLVLSLASISITKALLTRCPACSDNTFVNLCHTCTSPNOSL 135
DB 63 EPCPHLLTGDN-KLCTPSQABGLTKQAARHILGRCPSCFDNFALKWCEFTSPNOOD 121
QY 136 FINTVFAQL--GAGOLPAVAYEAF-----YQHS--PAEQSYDSCSRVRPAAATLAVG 186
DB 122 FVSISEMKPIEKGEGTPEYQAEAYVNTVEYRLSTDFAEGHFFSCKDVTFGGQPALAV- 180
QY 187 TWCVVYGSALCAQRAWLNFGQDTGNGL-APLDITFHLLEPGQAVSGIQP-LNEGVARCN 244
DB 181 -MC---TSPCTLTNWLFIQTQNLNIPHTKFLLYDPKTPPSDRSTYNNVNFVFTGCD 236
QY 245 ESGDDVATCSCODCAASCPAIPARQALDSTFYLQGM-----PGSLVLIILCSVPAAVT 299
DB 237 KSAVWGPACTSEC--NKEEYANLIDLDDGKTSQGTQCNVHGACINLFIWMLAFVLSLAV 294
QY 300 ILLVGF-----RVAPARDKSKMVDPKKGTSLSKLSFSFTLLLGQPF 341
DB 295 LLCVGFTSYSDYDYNLRQTQSGEESPRNRIK-----RTGAWI-----HNFENNA 342
QY 342 QCGTGWASWPLTILVLSVIVVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHF 401
DB 343 RDIGMWAGRNPKSHFFIGCAVLFLCPLGMIYHKESTNVVDMWSSPRRARQBEWVFANF 402
QY 402 GPFRFNQVLLTAPNRSSRYDLSLLGPKNFSGILDLLLELLELQERLRLHQLWSPEA 461
DB 403 GRPQRYQOIMLL--SHRDFOSSGKLYCP-----VFHKDIFEEFLFDILNAIKNISTQDSG 455
QY 462 QRNISLQICVAPLNPDTSLYDCCINSLLQYFQNNRTLLLTANOT-----L 509
DB 456 -RIITLDDVCYRMGPG-----YCLTMSPTNYEQGNKEHLDMSKNEETVSEDDDAFYF 510
QY 510 MGQTSQVDKWDHFLYCANAPLTPKQGTALALSCHADYGPVFPFLAIGGYKGDYSEAA 569
DB 511 SSEATTDEWNNHMAACIDQPMOK--TKSGLSCWGTGYPGSPAPNM-VFGKNSTNHQAANS 567
QY 570 LIMTFSLNYPAGDPRLAQAQKLEEAFLAEEMWRAFORRMAGMFOVFTTAERSLEDEINRTT 629
DB 568 IMMTILVTO--RTEPEIQKAELEKEKPLFKCKYREKSPKVI-FSFMARSIITDEIENDA 624
QY 630 AEDLPIFATSVIVIFLYISALGSY-----SSWSRVWSDSKATLGLGGVAVLGVMAAM 684
DB 625 KDEIVTVIALAFILGIVTFSLGRYFVCENQLWS--ILVHSR----- 664
QY 685 GFYSYLGISSLVLQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREVIHIGALGRVA 744
DB 665 GF----- 666
QY 745 PSMLLCSLSEACFFLGNALTPMPAVRTFALTSLGLAVLDFLLQWSAPVALLSDLSKROEA 804
DB 667 -----TDLPAIRTFCLYAGLAVLIDVVLHCTIFLALFVWDVQBELN 707
QY 805 SRLDVCCCKVQELPPPGQGEGL-----LLG-----FFQAYAPFLHLWITRGVV 849
DB 708 G-----KP-EFFFPFQIKDLGLAGYLIGRQATDTFTMQFFHFQVAPFLAHRMTIRIT 758
QY 850 LLLFLALFGVSLYSNCHISVGLDELALPKDSYLLDYFLFLNRYFEYFVGVAPYFVTTILGYN 909
DB 759 GIIFIASPITTVILSSKISVSGFDQSMAPTEKSYISTHFRYLDKDFDVGPPVFFVTVQGEGL 818

QY 910 FSSEAGMNAICSSAGCANNFSFTQKIQVATEFPEPQSYLAIPASSWVDQFIDWLT-PSCCCR 969
DB 819 WHPDVQNKFCITPGCSDTSTFGNIMYAVGHTETQYTLSGEMYNWIDNYLEWISRKSPCK 878
QY 969 LYISGPNKDFPCSTVNSL-----NCLKNCHSITWGSVRPSVEQPHKYLFWLNDRPNIKC 1024
DB 879 VYVHDPN--TFCSTNRNKSALDDKACRTCMDF----- 908
QY 1025 PKGLAAYSTSVNLTDGQVLASRFMAYHKL--KNSODYTEALRAAELAANITADLRK 1082
DB 909 --DGRASFKAISTSGRIQASQFMTHFKKLSISNSDFIKAMDYARMVSRRLERSI-- 964
QY 1083 VPGTDPAFEPFPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNL 1142
DB 965 ----DDTAHFVAFYSKIFPFYEQYSTINPILTTQFTVVGVGFIICVTGLGIDVKAACAV 1020
QY 1143 LSIYVILVDTVGFVALWDISYNAVSLNLNLSAVCMGVFVSHITRSEAIKPTWLERAK 1202
DB 1021 ICQVS-----NYFHIWSS--GILIEFSVNLKGYACSLURAKORAE 1060
QY 1203 EATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLITLLGLHLGLVFLPVIL 1262
DB 1061 STVSGIPIILSGPVTVWAGSTMFLSGAHLQIITVYFFKLFLITIVSSAVHALIILPILL 1120
QY 1263 SYVGPDVNPALALBQKRAEAAVAAVMVASCNHPSPRVSTADNIYV-----HSPEGS 1314
DB 1121 AFGSRGHGSSETSTNDNDQHDACVLS--PTABSHISNVBEGILNRPSLDASHILDPL 1178
QY 1315 IKGAGATSNFLPNNGRQF 1332
DB 1179 LKAEGLIDKAI---GRDF 1193
RESULT 28
AAM80153
ID AAM80153 standard; protein; 229 AA.
XX
AC AAM80153;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 3799.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00663561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK53286.

PF 22-MAR-2001; 2001WO-JP002279.
PR 24-MAR-2000; 2000JP-00088595.
XX (TAKE) TAKEDA CHEM IND LTD.
XX Taniyama Y, Kita S, Komiyama T;
XX WPI; 2001-611501/70.
XX N-PSDB; AAI66914, AAI66931.
XX
XX New sterol-sensing domain-containing protein for diagnosing and screening
XX candidate compounds in drug development for diabetes, obesity, cancer,
XX arteriosclerosis, hyperlipidemia and neurodegenerative disorders.
XX
XX Claim 6; Page 113-114; 171pp; Japanese.
XX
XX The invention provides a novel SSD (sterol-sensing domain)-containing
XX protein. The protein originates from human liver, human testis or human
XX brain. The protein can be expressed by standard recombinant methodology.
XX The proteins, encoded DNAs and antibodies are useful in diagnosis and
XX screening candidate compounds in drug development for diabetes, obesity,
XX cancer, arteriosclerosis, hyperlipemia, neurodegenerative disorders such
XX as Alzheimer's disease and neural disorders. The present sequence
XX represents a human SSP1 SSD domain
XX
XX Sequence 194 AA;
SQ
Query Match 13.6%; Score 942; DB 4; Length 194;
Best Local Similarity 99.5%; Pred. No. 5e-79; Mismatches 1; Indels 0; Gaps 0;
Matches 193; Conservative 0;
QY 615 FTAEKSLDEINRTAEDLPFATSVIVIFLYISALGYSWSRWVDSKATLGLGVA 674
DB 1 FWAERLEDEINRTAEDLPFATSVIVIFLYISALGYSWSRWVDSKATLGLGVA 60
QY 675 VVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREV 734
DB 61 VVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREV 120
QY 735 HIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTFALTSGLAVIDFLQMSAFVAL 794
DB 121 HIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTFALTSGLAVIDFLQMSAFVAL 180
QY 795 LSLDSKRQASRLD 808
DB 181 LSLDSKRQASRLD 194
RESULT 31
AAR75375
ID AAR75375 standard; protein; 1447 AA.
XX
AC AAR75375;
XX
XX 30-JUL-1996 (first entry)
XX
XX Human patched protein.
XX
XX Patched gene; PTC protein; embryo development; cellular regulation;
XX signal transduction; ligand; antibody; hedgehog protein.
XX
XX Homo sapiens.
XX
XX WO9611260-A1.
XX
XX 18-APR-1996.
XX
XX 06-OCT-1995; 95WO-US013233.
XX
XX 07-OCT-1994; 94US-00319745.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.

XX Scott MP, Goodrich LV, Johnson RL;
XX WPI; 1996-209842/21.
XX N-PSDB; AAT14220.
XX
XX DNA encoding patched protein other than Drosophila patched protein - used
XX to produce antibodies which detect or inhibit patched protein ligand
XX signal transduction in cells.
XX
XX Example; Page 56-61; 70pp; English.
XX
XX Human patched protein (PTC) (AAR75375) was identified as the product of a
XX cDNA clone (AAT14220) derived from human lung. It has 96% identity and
XX 98% similarity to mouse PTC. PTC protein has been proposed as a receptor
XX for hedgehog protein on the basis of genetic experiments in flies. Human
XX PTC protein can be obtained in large amounts by expression of the cDNA clone
XX in transformed host cells. It can be used to screen for agonists and
XX antagonists to isolate its ligand, partic. Sonic hedgehog, to assay for
XX the transcription of ptc mRNA and to raise antibodies
XX
XX Sequence 1447 AA;
SQ
Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 2.4e-50;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PGQAVSGIQFLNEGVARCNESQGDVATCSQDCAACPAIARPQALDSTFVLGMPGS 284
DB 24 PGRAGGRRRTGGLRAAAPDRD-----YLRPSYCDAAFALEQI--- 65
QY 285 LVLIILLCSPAVVTILLVGRVAPARDKSKMDPK-----KGTSLSKLSFSTHTLL 337
DB 66 -----SKGATGRKAPLWRAKFORLLFKGCVIQNC 98
QY 338 GQFFQGWGTWASWPLTILVSLVIVVALAAGLVFTETDTPVELWSAPNSQARSEKAFH 397
DB 99 GKX-----LVVGLLIFGAFVGLKAANLETNVEBELVVEGVRVRELNYT 143
QY 398 DQHFGPFRFTY-QVILTAPNRSSVRYDLSLLGPKNFSGIILDLLELLE--LQERLRLH 454
DB 144 RQKIGEEAMFPOLMIQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QVMSPEAQRNLSQDLCVAP-----LMPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 191 YMYN-----RQWLEHLCKYSGELITETGYMQLIIEYLVPCLIITPLDCFWEGAKLQSGTA 246
QY 506 NQTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 --YLLGKPLRWNTNFDPLFLEELKKNYQVDSWEEMLNKAEGVGHYMDRPLCLNPADPDC 304
QY 528 -----APL-----TFKDG-----ALALSCWA 544
DB 305 PATAPNKNSTKPLDMALVNGGCHLSRKYHWQBELIVGGTVKNTCKLYSAHALQTMF 364
QY 545 DYGAFFVFPFLATGKGYKHYSAAEALIMTFSLNYPAGDPPLRAQAKLNEEAFLEMPAPQ 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAAILLEAQRTVVEVHOSV 411
QY 605 RMAGMFQVTPFAERSLEDEINRTAEDLPFATSVIVIFLYISALGYSWSRWVDS 664
DB 412 AQNSTQKVLSEFTT-TLDDILKSPSDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGGVAVVLGAVMAAMGFSSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRL 724
DB 466 QGAVGLAGVLLVALSVAGLGLCSLIGISFNAAATTVQLVFLFALGVGDDVFLAHAFSET 525
QY 725 PRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTFALTSGLAVIDFL 784
DB 526 GQNKRIFFEDRTGECLKRTGASVALTSISNVTAFFMAALIPFALRAFSLQAAVVVVFNF 585
QY 785 LLOMSAFVALLSLDSKRQASRLDVCCC-----VKQEL-----PPPG 822

Db	586	AMVLLIPAILSMYDLYRREDRLDIFCCFTSPCVSRVIOVEPQAYTDTHNTRYSPPPY	645
Qy	823	QCEGL-----	827
Db	645	SSHSPAHEQTITMQSTVQLRTEYDPHTVHYVYTAEPSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFPQKAYAPFLLHMTIRGVVLLLFALPGVSL	861
Db	706	SSSTRDLISQFSDSLHCLPEPPCTKWTLSSPAERKHYAPFLKPKAKVWVIFLPLGLIGVSL	765
Qy	862	YSMCHISVGLDQELAPKDSYLLDYFLNRRFEVGCAPVYFVTLGYNFSSSEGNMAICS	921
Db	766	YGTTRVRDGLDLTDIVPRTREYDFIAQFKYFSF-----YNN-----	803
Qy	922	SAGCNNFSFTQKIYAT-----EPPEQSYLAIPASS-----WVDDFDIML-----	961
Db	804	-----YIVTORADYENIQHLLYDLHRSFNKYVMLEENKQLPKWMLHYFRDMLQGLQD	857
Qy	962	-----TPSS-----CCRLYISGPNKDKFCPSTVNSLNCNKMSITMGS	1000
Db	858	AFDSDWETGKIMPNNYKNGSDDGVLAYKLLVOTGSRDK--PIDISQLT--KQRLVDADGI	913
Qy	1001	VRPSVEQPHYL--PWFLND-----RPNKCPKGLAAY--STSNNLTSDQ	1043
Db	914	INPSA--FYIYLTAWNSNDPVAYASQANIRPHRPEVHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASRFMAYHKPKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFVFPVYTTITNVFY	1102
Db	969	IEYAQPPFYGLNGLRDTDFEALKEKVTTCNSYTSGLSSYPNG-----YFF-----LFW	1018
Qy	1103	QYVLTLPGLPMLSLCVPTFAVSCLLGLDLRSLGLMLLSIVMLVDTVGVFMALWDIS	1162
Db	1019	EQVIGLRHMLLLFISVLACTELVCVAVFLNFWTAGII--VMVLAIMTVBELFGMWGLIGK	1077
Qy	1163	YNVSLINLVSAVGSVEFVSHITRSP--AISTKPTWLERAKERTISMGSAPFAGVAMTN	1220
Db	1078	LSAVPVVILLIASVGIGVEFVTHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
Qy	1221	LPGIIVLGLAKAQIQIPFRLMLLITLGLLHGLVFLPVILSYVG--PDVNPALALEQ-	1277
Db	1134	LGVLMLAGSEDFIVRYFVAVLAILTILGVNLGLVLLPVLLSPFGPYEVSFANGLRNL	1193
Qy	1278	--KRAEEAVALVAVVACPNHPSRVSTADNIYVNHSPGSIK	1317
Db	1194	PTPSEPPPSVVRFAMPFGHTH--SGSDSSDSEYSSQTTVSG	1233
RESULT 32			
ID	AAW52200		
XX	AAW52200 standard; protein; 1447 AA.		
AC	AAW52200;		
XX	25-JUN-1998 (first entry)		
DE	Human patched (ptc) protein.		
KW	Patched protein; ptc; cancer; tumour suppressor; cell adhesion promoter;		
KW	wound healing; ageing; human.		
OS	Homo sapiens.		
XX	WO9745541-A2.		
XX	04-DEC-1997.		
XX	02-JUN-1997; 97MO-US009553.		
XX	31-MAY-1996; 96US-00656055.		
PA	(STRD) UNIV LELAND STANFORD JUNIOR.		
PA	(REGC) UNIV CALIFORNIA.		
XX			

PI	Scott MP, Goodrich LV, Johnson RL, Epstein E, Oro A;
XX	WPI: 1998-032648/03.
DR	N-PSDB; AAV21590.
XX	Patched protein other than Drosophila melanogaster patched protein - used for characterising the phenotype of a tumour.
PT	Claim 4; Page 76-80; 86pp; English.
XX	This is a human patched (ptc) protein. The encoding DNA can be used to construct an expression cassette comprising an altered patch or hedgehog gene. The expression cassette comprises a nucleic acid encoding a patched protein other than a Drosophila melanogaster patched protein, or fragment of at least 12 nucleotides in length, as other than an intact chromosome under transcriptional control of a transcriptional initiation region, and a transcriptional termination region, both functional in an expression host. A genetically engineered mammalian cell comprising this expression cassette as an extrachromosomal element or integrated into the genome of the cell can be predisposed to develop basal cell carcinoma as a result of the transfection. By analysing DNA, functional analysis of patched protein function, or by detecting antibody binding to abnormal patched protein, a genetic predisposition to developmental abnormalities and cancer can be diagnosed. This analysis can also be used for characterising the phenotype of a tumour, particularly a carcinoma, especially a basal cell carcinoma. The methods can also be used for characterising transitional cell carcinoma of the bladder, meningiomas, medulloblastomas, etc. The modified cells comprising the expression cassette can be used to determine the role of different exons of the patched gene in oncogenesis, signal transduction, etc. Transgenic animal models created from these cells can be used as animal models for carcinomas of the skin. The patched protein of mosquito, butterfly or beetle or alternatively, a mammalian patched protein of human or mouse can be used to identify ligands or substrates that bind to, modulate, or mimic the action of patched gene. These agents could be used as tumour suppressors, cell adhesion promoters (e.g. in wound healing and ageing)
XX	Sequence 1447 AA;
QY	Query Match
DB	Best Local Similarity 9.4%; Score 651.5; DB 2; Length 1447;
QY	Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY	225 PGQAVSGIQPLNEGVARCNESQGDVATCSCQDCAACPAIARQALDSTFVLGMPGS 284
DB	24 PCRAGGRRRTTGLRAAADPRD-----YLRPSYCDRAFALEGI--- 65
QY	285 LVLIIILCSVFAVWTTILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
DB	66 -----SKGATGRKAPLMRAKFORLLFKLGCYIQKNC 98
QY	338 QGFFQGWGTWASWPLTILVSVIPVALAAGLVFTELTDPVELWSAPNSQARSEKAPH 397
DB	99 GKF-----LVVGLLIIFGAFVGLKAANLTNNVELVBEVGGVRSBELNYT 143
QY	398 DOHGFPPFRTN-QVILTAPNRSSVRYDSLLGPKNFSGIILDLLELLELLE--LQERLRHL 454
DB	144 RKIGEEAMFAPQLMIQTPEEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY	455 QVMSPEAQRNLSLDICYP-----LNPNTSLDYCCINSLLQYFQNNRTLLLLTA 505
DB	191 YMYN---ROWKLEHLCYKSGELITETGYMDQIIEYLYPCLITPLDCFWEGAKLSGTA 246
QY	506 NOTLMGQ-----TSQVDKDHFLY-----CAN----- 527
DB	247 ---YLLGKPLRWTFNPDPLEFLEELKKINYQVDSWEEMLNKAIFVGHGYMDRPCINPADPC 304
QY	528 -----APL-----TFKDG-----ALALSQWA 544
DB	305 PATAPNKNSTKPLDMALVNLGCGHLSRKYMHQBELIVGGTVKNSGKLVSAHALQTMF 364
QY	545 DYGAVPVFPFLAIGGYGKHGYSEAEALIMTFISLNNYPAGDPRLAQAQKLWEEAFLEEMRAFQ 604

365	QLMTPKQWY----	EHPKGYEY-----	VSHINWNE-----	DKAAAILAEAWQRTYVEVHQSV	411	KW	Human; patched gene; diagnosis; treatment; developmental disorder;
605	RMAGMFOVTTAERSLEINRTAEDLPATSYIVIFLYISIALGSSYSSRWVMDS	664	XX	XX	XX	KW	cancer; healing; injured tissue; ptc; spina bifida; Wnt-1 oncogene;
412	AQNSTQKVLSTTT--TTLDDILKSFSDSVIRVASGYLLMLAYACLTW---LRWD--CSKS	465	OS	XX	XX	OS	sperm production; gene therapy.
665	KATGLGVAVVVLGVAMAMGFSSYLGIRSLVLQVVPFLVLSVGADNPIFVLEVQRL	724	PN	XX	XX	PN	Homo sapiens.
466	QCVAGLAVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGVDDVPELLAHAFSET	525	XX	XX	XX	XX	US5837538-A.
725	PRRQEPREVIHGRALGRVAPSMILCSLSAICPFLGALTTPMAVTRFALTSLGLAVILDF	784	XX	XX	XX	XX	17-NOV-1998.
526	GQNKRIPEFDRTEGCKRTGASVALTSISNTVAFMAALIPIPALRAFSLQAAVVVWFNF	585	XX	XX	XX	XX	06-OCT-1995; 95US-00540406.
785	LLQMSAFVALLSLDSKROEASRLDVCCC-----VKPQEL-----	822	XX	XX	XX	XX	07-OCT-1994; 94US-00319745.
586	ANVLLIFPAILSLMDLYRRREDRLDIFCCFTSPCVSRVLOVEPQAYTDHNTRYSPPPPY	645	XX	XX	XX	XX	(STRD) UNIV LELAND STANFORD JUNIOR.
823	OQEGLE-----	827	XX	XX	XX	XX	Scott MP, Goodrich LV, Johnson RL;
646	SSHSPAHETQITMQSTVQLTRTEYDPHTHYVYTTABPRSEISVQPVTVTQDTLSCQSPST	705	XX	XX	XX	XX	WPI; 1999-023461/02.
828	-----LLGFFQKAYAPFLHWTIRGVVLLLFALFGVSL	861	XX	XX	XX	XX	N-PSDB; AAV64093.
706	SSTRDLLSQFSDSLHCLPEPCTKWTLSFSAEKYAPFLPKAKVWVIFLGLLGVSL	765	XX	XX	XX	XX	Nucleic acid encoding vertebrate patched protein and related
862	YSMCHISVGLQELALPKDSYLLDYFLNRYFEVGAPVYFVTTLGYNFSSEAGMNAICS	921	XX	XX	XX	XX	transformants - used to express poly(peptide(s), useful for diagnosis and
766	YGTRVRDGLDITDVPRETREYDFAAQFYFSP-----YNN-----	803	XX	XX	XX	XX	treatment of developmental disorders or cancer, and in healing of injured
922	SAGCNFFSTQKIQVAT-----BPPEQSYLAIPASS-----WDDDFIDWL-----	961	XX	XX	XX	XX	tissue.
804	-----YIVTQADYPNIQHLLYDLHRSFSNRYVWLEENKQLPKWHLHYFDMLQGLQD	857	XX	XX	XX	XX	Claim 23; Col 63-70; 38pp; English.
962	-----TPSS-----CCRLYISGPNKDKFCPSTVNSLNCNCSITMGS	1000	XX	XX	XX	XX	The present sequence represents the human patched (ptc) protein. Cells
858	AFSDMETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913	XX	XX	XX	XX	containing and expressing the ptc gene are used for the recombinant
1001	VRPSVEQFKHYL--PWFLND-----RPNIKCPKGLAAY--STSVNLTSDGQ	1043	XX	XX	XX	XX	production of the protein. These in turn are useful: (i) for generating
914	INPSA--FYIYLTAWVSNDPVAVAASQANIRPHRPEWVHDK--ADYMBETRLRIPAAEP	968	XX	XX	XX	XX	antibodies (Ab); and (ii) to screen for specific-binding ligands
1044	VLASRFMAYHKPLKNSQDYTEALRAARELANITA--DLRKVPCTDPAFVFPVTTNVFY	1102	XX	XX	XX	XX	(potential therapeutic agonists and antagonists). The ptc gene, or its
969	IEYAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSLSGLSSYPNG-----YFP-----LFW	1018	XX	XX	XX	XX	fragments, are used to isolate related sequences from other mammals; to
1103	EQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLGMLNLSIVMLVDTGVGFMALWDIS	1162	XX	XX	XX	XX	identify mutations (particularly those associated with genetic diseases
1019	EQYIGLHWLLLFISVVLACTFLVCVFLNPNWTAGII--VMVLAMTVLFGMMGLIGIK	1077	XX	XX	XX	XX	such as spina bifida and other developmental disorders); to monitor
1163	YNAVSLINLVSAGVMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSVAPGAVAMTN	1220	XX	XX	XX	XX	expression levels in testis (to determine relationship with sperm
1078	LSAVPVVILLIASVGIGVEFTVHVALAFLTAIGDKN---RRAVLALSHMFAPVLDG-AVST	1133	XX	XX	XX	XX	production) and to isolate 5'-non-coding sequences (used to study
1221	LPGLVLGLAKAQLIQIFPRFLMILLITLGLLHGLVPLPVLSVVG--PDVNPALALEQ-	1277	XX	XX	XX	XX	embryonic development and to provide regulated expression of proteins).
1134	LLGVMLAGSEFDFIVRYFVAVIAILITLGVNLGLVLLPVLLSFFGYPYEVSPANGNLRL	1193	XX	XX	XX	XX	The complete gene can be used in gene therapy, including expression of
1278	--KRAEEAAMVAVASCPNHPSRVSTADNIYVNHSPFSGSIKG	1317	XX	XX	XX	XX	antisense molecules, and to generate transgenic animals for studies of
1194	PTSPPEPPSPVVRPAMPFGHTH--SGSDSDSDSEYSSQTTVSG	1233	XX	XX	XX	XX	protein on cell surfaces and as competitive inhibitors to determine the ptc
225	PGQAVGSGIQPLNEGVARNESQGDVATCQDCAASCPATARQALDSTFVLQMPGS	284	QY	225	PGQAVGSGIQPLNEGVARNESQGDVATCQDCAASCPATARQALDSTFVLQMPGS	284	Query Match
24	PCRAGGGRRRRTTGLRAAAPDRD-----YHRRSYCDAAFALEQI---	65	Db	24	PCRAGGGRRRRTTGLRAAAPDRD-----YHRRSYCDAAFALEQI---	65	Best Local Similarity 9.4%; Score 651.5; DB 2; Length 1447;
285	LVLIIILCSVFAVWVILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSFHTLL	337	QY	285	LVLIIILCSVFAVWVILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSFHTLL	337	Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
66	-----SKGATGRKAPLWLRKAFQRLFKLGCYKQNC	98	Db	66	-----SKGATGRKAPLWLRKAFQRLFKLGCYKQNC	98	
338	GOFFQCGWTWASWPLTILVSVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKATH	397	QY	338	GOFFQCGWTWASWPLTILVSVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKATH	397	
99	GKPF-----LVVGLLIFGAFVAGLKAANLETNVEELVVEGVGRVSRRELYNT	143	Db	99	GKPF-----LVVGLLIFGAFVAGLKAANLETNVEELVVEGVGRVSRRELYNT	143	
398	DQHFGFFFRTN-QVILTAPNRSYRYDLSLLGPKFNKSGILDLDLLELLE--LQERLRLH	454	QY	398	DQHFGFFFRTN-QVILTAPNRSYRYDLSLLGPKFNKSGILDLDLLELLE--LQERLRLH	454	
144	RQKIGBEAMFNPMQIMQTPKEEG-----ANVLTTEALLQHLDSALQASRVHV	190	Db	144	RQKIGBEAMFNPMQIMQTPKEEG-----ANVLTTEALLQHLDSALQASRVHV	190	
455	QVWSPQAQNISLQDICYAP-----LNPDNTSLYDCINSLLOYFQNNRTLLLLTA	505	QY	455	QVWSPQAQNISLQDICYAP-----LNPDNTSLYDCINSLLOYFQNNRTLLLLTA	505	
RESULT 33							
AAW72969							
ID	AAW72969						standard; protein; 1447 AA.
XX	AAW72969						
XX	AAW72969						
XX	26-JAN-1999 (first entry)						
DT	Human patched protein.						
DE							
XX							

Db 191 YMYN-----ROWKLEHLCYKSGELITETGYMDQIIIEYLYPCLLIITPLDCFWEGAKLQSGTA 246
Qy 506 NOTLMGO-----TSQVDMKDHFLY-----CAN----- 527
Db 247 --YLLGKPELWNTWDFPLEBLELKKINYQVDSHEMLNKAELVGHGYMDRECLNPADPDC 304
Qy 528 -----APL-----TFXQGT-----ALALSCMA 544
Db 305 PATAPNKSTKPLDMALVNGCHGLSRKYMHWOBELIVGCTVKNSTGKLVSAHALQTMF 364
Qy 545 DYGAIPVFPFLAIGGYKGYDSEAEALINTFSINNYVAGDPRLAQAKLWEBAFLSEMAFQ 604
Db 365 QLMTPKQMY---EHPKGYEY-----VSHINWNE---DKAAALEAWQRYVEVHQSV 411
Qy 605 RRMAGMFOVTFPAERSLEDEINRTTAEDLPFATSYIVIFYISIALGYSYSSRWVMVDS 664
Db 412 AQNSTQKVLSTFT-TTLDDILKSFSDSVIRVASGYLLMLAYACLTM--LRWD--CSKS 465
Qy 665 KATLGLGVAVVGLVAMAMGPFYSVLGRSSLVILQVVPFLVLSVGADNIFIFVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLFFLALGVGVDDVFLLAHFSET 525
Qy 725 PRPGEPREVIHGRALGRVAPSMILCSLSEALCFELGALTMPAVRTEALTSGLAVIDF 784
Db 526 GQNKRIPEEDRTGECUKRTGASVALTSISNVTAFPMALIPALRAEFLQAQAVVVFNP 585
Qy 785 LQMSAFVALLSLDKROBASRLDVCCC-----VKPQEL-----PPPG 822
Db 586 AMVLLIFPALLSMDLYRREDRLDIFCCPTSPCVSRVQVEPQAVTDTHDNTRYSPPPY 645
Qy 823 QSEGL----- 827
Db 646 SGHSFAHEQITMQSTVQLRTEYDPHTHYVYTTAEPRSEISVQPVVTQDTLSCOSPST 705
Qy 828 -----LLGFFOKAVAPFLHWTIRGVLLLLFLALFGVSL 861
Db 706 SSTRDLLSFSSSLHCLBPCTKWTLSFSAEKHAPFLPKPKAKVVIFLFLGLLQVSL 765
Qy 862 YSMCHISVGLDQELAPKDSYLLDYFLFLNRYFVGVGAPYVFTTILGYNFSSBAGNAICS 921
Db 766 YGTRVRDGLDLTDIVPRETREYDPTAAQKYSF-----YNN----- 803
Qy 922 SAGCNFPSTQKIQYAT-----EPQESYLAIPASS-----WVDDFIDML----- 961
Db 804 -----YIVTQKADYFNIQHLLYDLHRSFNKYMLENKQLPKMWHYFRDMLQGLQD 857
Qy 962 -----TPSS-----CRLYISGNKDKFCPTSVNSINCLKNCSITMGS 1000
Db 858 AFSDWETGKIMPNYKNGSDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGI 913
Qy 1001 VRPSVEQFHKYL--PWFLND-----RPNIKCPKGGLAAY--STSNNLTSDGQ 1043
Db 914 INFSA--FYIYLTAWVSNDDPVAASAQANIRPHRPEWVHDK---ADYMPETRLRIPAAEP 968
Qy 1044 VLASRFMAYHKPLKNSDYTEALRABELAANITA-DLRKVPCDTPAFEPFYTITNVFY 1102
Db 969 IEVAQFPFYLNGLRDTSDFEVIAEKVRTICSNYTSLSGLSSYPNG-----YFP-----LFW 1018
Qy 1103 EOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLLSIYMLIVDTVGFPMALWDIS 1162
Db 1019 EOYIGLRHLLWLLFISVVACTFLVCVFLNPNPTAGII--VMVLAMTVELFGMWGLIGIK 1077
Qy 1163 YNAVSLINLVSAGMSYFVSHITRSP--AISTKPTWLERAKAATTSMGSAVAGVAMTN 1220
Db 1073 LSAVPVVLIIASVGIQVEFTVHVALAFLTAIGDKN---RRAVLALHEMFAPVLIDG-AVST 1133
Qy 1221 LPLGLVLGLAXAQIIQIFFRALNLLITLLGLLHGLVFLPVLTSYVG--PDVNPALALEQ- 1277
Db 1134 LLGLVLMAGSFDPIFVIRYFAVLAITLILGVNLGLVLLPILLSFPGYPPEVSPANGLNRL 1193
Qy 1278 --KRAEBAVAAVMVASCPNHPRSRVSTADNIYVNHSPFEGSING 1317
Db 1194 PTPSPPEPPSVVRPAMPFGHTH--SGSDSSDSEYSSQTTVSG 1233

RESULT 35
AAE19830
ID AAE19830 standard; protein; 1447 AA.
XX
AC AAE19830;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human patched (Ptc) protein.
XX
KW Human; patched; Ptc; cell proliferation; differentiation; therapy;
KW cyostatic; testicular cancer; hedgehog protein signalling.
OS Homo sapiens.
XX
FH Key
FT Domain
FT
FT /label= TM1
FT /note= "Transmembrane domain"
FT
FT /label= TM2
FT /note= "Transmembrane domain"
FT /label= TM3
FT /note= "Transmembrane domain"
FT /label= TM4
FT /note= "Transmembrane domain"
FT /label= TM5
FT /note= "Transmembrane domain"
FT /label= TM6
FT /note= "Transmembrane domain"
FT /label= TM7
FT /note= "Transmembrane domain"
FT /label= TM8
FT /note= "Transmembrane domain"
FT /label= TM9
FT /note= "Transmembrane domain"
FT /label= TM10
FT /note= "Transmembrane domain"
FT /label= TM11
FT /label= TM12
FT /note= "Transmembrane domain"
XX
XX US6348575-B1.
XX
XX 19-FEB-2002.
XX
XX 15-APR-1999; 99US-002939505.
XX
XX 15-APR-1998; 98US-0081884P.
XX
XX (GETH) GENENTECH INC.
XX
XX De Sauvage P, Carpenter DA;
XX
XX WPI; 2002-215260/27.
XX
XX Native human patched-2 polypeptide for treating disorders caused by
XX hedgehog protein signaling such as testicular cancer, and for screening
XX cDNA libraries.
XX

Example 1; Fig 3; 82pp; English.

PS The invention relates to an isolated sequence comprising a native human
 CC patched2 (Ptc2) polypeptide. The invention also relates to signalling
 CC and mediator molecules in the hedgehog (hh) cascade which are involved in
 CC cell proliferation and differentiation. The isolated sequence is useful
 CC for the treatment of disorders which are linked to Hedgehog, especially
 CC Desert hedgehog expression, such as testicular cancer. It may also be
 CC used as a hybridisation probe in a cDNA library to isolate Ptc2 or its
 CC homologues, and to diagnose whether a disorder is driven by Ptc2 or
 CC Hedgehog protein signalling. The present sequence is human patched (Ptc2)
 CC protein. Note: The sequence shown as SEQ ID NO: 4 in sequence listing of
 CC the specification is a fragment of the sequence shown as SEQ ID NO: 4 in
 CC fig 3 of the specification

XX Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 5; Length 1447;

Best Local Similarity 22.0%; Pred. No. 2.4e-50;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGAGVSGIQLNEGVARNESQGDVATCSQDCAASCPAIAAPQALDSTFVLGMPGS 284

DB 24 PEPAGGRRRTGRLRAAPDRD-----YLRPSYCDAAFALEQI--- 65

QY 285 LVLIILCSFVAVVITLLVGFVRVAPARDKSKMDPK-----KGTSLSDKLSFSFTHLL 337

DB 66 -----SKGKATGRKAPLWLRAKQFRLFLKGCYIQKNC 98

QY 338 GQFFQGWGTWASWPTLILSVIPVVALAAGLVFTELTDVPELMSAPNSQARSEKAFH 397

DB 99 GRF-----LVGGLIFGAPAVGLKAANLTNEVELWVEVGRVSRLENTY 143

QY 398 DQHFQPPFRTN-QVILTAENRSSRYDLSLLGPKNFSGLTDLDDLELE--LQERLRL 454

DB 144 RQKIGEAENFQMIQTPEEG-----ANVLTEALLQHLDSALQASRVHV 190

QY 455 QVMSPEAQNIQLDICYAP-----LNPDTSLYDCCINSLLIQYQNNRTLLLLTA 505

DB 191 YMYN---RQWLEHLCKYSGELITETGYMDQIIEVLYPCLITPLDCFWGAKLQSGTA 246

QY 506 NOTLMGQ-----TSQVMDKHFLY-----CAN----- 527

DB 247 ---YLLGKPLRWTFDPLBFLBELKKINYQVDSWEEMLNKAEBVGHGYMDRPLNADPDC 304

QY 528 -----APL-----TFKQGT-----ALALSNA 544

DB 305 PATAPNKNSTKPLDMLVNGCHGLSRKYMWQBELIVGGTVKNSTGKLVSAHALQTMF 364

QY 545 DYCAPVFPFLAIGYKGYDSEAEALIMTFSLNYPAGDPRLAQAQKMBEAPLEENRAPQ 604

DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE---DKAAALLENQRTYVEVHVQSV 411

QY 605 RMAGHFQVTPFAERSLDEINRTTAEIDLPIFATSYIVIFLYISIALGYSYSSWRVMWDS 664

DB 412 AQNSTOKVLSFTT-ITLDDILKSFSDSVIRVASGYLLMLAVACLTM---LRWD---CSKS 465

QY 665 KATLGGVAVVLGAWMAWGFYSYLGIRSSVLQVPPVLSVGCADNIFIFVLEYQRL 724

DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGTSFNAAATQVLPFLALGVGVDDVFLAHAFSET 525

QY 725 PRRPGPREVHIGRALGRVAPSMMLCSLSEALCFGLGALTMPVAVRTPALTSGLAVILDF 784

DB 526 GQNKRIFFEDRTEGCLKRTGASVALTSISNVTAFMAALIPALBAFSLQAQVAVVNF 585

QY 785 LLOMSAFVALLSDSKRQASRLDVCCC-----VKQDEL-----PPPG 822

DB 586 AMVLIFPAILSMDLVRRDRDLDFCCPTSPCVSRVIQVEFQAVTDTHTNTRYSPPPY 645

QY 823 QEGL----- 827

DB 646 SSHSAFHEQTITMSTVQLRTEYDPHTHYVYTTAEPRSEISVQPVTVTQDTLSCOSPEST 705

QY 828 -----LLGFFQKAYAPFLLHWTITRGVLLFLALFGVSL 861

DB 706 SSTDLLSQFSSSLHCLPEPPCTKWTLSFAEKHYAPFLLKPKAKVWVIFLFLGLGVSL 765

QY 862 YSMCHISVGLDDELALPKDSYLLDFLNRFFEVGAPVYFTTLCYNSSESGMNAICS 921

DB 766 YGTTVRDGLDITDVPRETRYDFTAAQKYPSP-----YNN----- 803

QY 922 SAGCNFSTQKIQYAT-----EPQOSYLAIPASS-----WVDDFIDWL----- 961

DB 804 -----YVYQKADYPTQHLIDYLRFSFNKYVWLEENKQLPKMWLHYFRDQLQLOD 857

QY 962 -----TPSS-----CCRLYISGPNKCKPCPSTVSNLCLNCKMCSITMGS 1000

DB 858 AFDSDWETGKIMPNNYKNGSDDGVLAYKLLVQVGRDK--PIDISQLT--KQRLVDADGI 913

QY 1001 VPSPVQPHKYL-PWFLND-----RPNICKPKGLAAY--STSVNLTSDDQ 1043

DB 914 INPSA--FYIYLTAWVSNDFVAYAAASQANIRPHRPFWHDK---ADYMPETRLRIPAAEP 968

QY 1044 VLASRFMAYHKPKNSQDYTEALRAARELAANITA-DLRKVPCTDPAPFVFPVTTINVFY 1102

DB 969 IEYAQPPFVYNGRLDTSDFEAEIKVRTICSNYTSGLSSYPNG-----YDF-----LFW 1018

QY 1103 EQLYTLPEGLPMLSLVPTFAVSCLLGLDLRLSGLMLLSIVMLVDTVGFMAWDITS 1162

DB 1019 EQYIGLRHLLLFISVLACTFLVCVAFLLNPWTAGII-VMVLAALMTVELFGMGLIGIK 1077

QY 1163 YNAVSLINLVASGVMSVRFVSHITSPF--AISTKPTWLERAKENTISMGSAVPAGVAMTN 1220

DB 1078 LSAPVWVILIASVGIVGEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133

QY 1221 LPGIILVLGAKAQLIQIPFFRLNLITLGLLHGLVFLPVILSYVG--PDVNPALALEQ- 1277

DB 1134 LGVLMLAGSEDFVIRVFVFAVLAITILGVNLGLVLLPVLLSFFGYPFVSPANGNLRL 1193

QY 1278 --KRAEAAVAAVMVASCNHPHSRVSTADNIYVNHSPFSGIK 1317

DB 1194 PTPSPPPSVVRFAMPFGH--SGSDSDSEYSSQTTVSG 1233

RESULT 36

ABJ10931

ID ABJ10931 standard; protein; 1447 AA.

XX

AC ABJ10931;

XX

DT 12-DEC-2002 (first entry)

XX

DE TRC8 related human patched protein SEQ ID No 5.

XX

KW TRC8; Translocation in Renal cancer from Chromosome 8; fused DNA; 3.2;

KW FHIT/TRC8 fusion DNA; sporadic renal cell carcinoma; TRC8/FHIT; TRC8FHIT;

XX

OS Homo sapiens.

XX

PN US2002106656-A1.

XX

PD 08-AUG-2002.

XX

PF 02-JUL-2001; 2001US-00898533.

XX

PR 12-MAR-1998; 98US-0077723P.

PR 12-MAR-1999; 99US-00268140.

XX

PA (GENM/) GENMILL R M.

XX

PI (DRAB/) DRABIN H A.

XX

PI Gemmill RM, Drabkin HA;

XX

XX WPI; 2002-712395/77.

XX

[illegible]

XX	Human patched gene PTC product.	
DE	Human; patched gene; PTC; hedgehog protein; gene therapy.	
XX	Homo sapiens.	
OS	US2003032085-A1.	
XX	13-FEB-2003.	
XX	20-OCT-1997; 97US-00954701.	
XX	07-OCT-1994; 94US-00319745.	
PR	06-OCT-1995; 95US-00540406.	
XX	(SCOTT/) SCOTT M P.	
PA	(GOOD/) GOODRICH L V.	
PA	(JOHN/) JOHNSON R L.	
XX	Scott MP, Goodrich LV, Johnson RL;	
PI	WPI; 2003-492065/46.	
DR	New DNA sequence other than present in a chromosome encoding patched gene	
XX	other than Drosophila patched gene, useful for preparing transgenic	
PT	laboratory animals and to knock out patched protein in embryonic stem	
PT	cells.	
XX	Disclosure; Page 8-10; 40pp; English.	
PS	The invention relates to a DNA sequence other than present in chromosome	
XX	encoding a patched (ptc) gene other than Drosophila patched gene. A cell	
CC	expressing the DNA sequence is useful for producing patched protein, by	
CC	growing the cell expressing the DNA sequence, where the patched protein	
CC	is expressed and isolating the patched protein free of other proteins.	
CC	The cell expressing the DNA sequence is also useful for screening	
CC	candidate compounds for binding affinity to the patched protein, by	
CC	combining the candidate protein with the cell expressing the DNA sequence	
CC	where the DNA sequence comprises the entire coding sequence under the	
CC	transcriptional regulation of the transcriptional initiation region and a	
CC	transcriptional regulation region functional in the cell, expressing the	
CC	patched protein in the cell and assaying for the binding of the candidate	
CC	compound to the patched protein. The above cell is useful for screening	
CC	candidate compounds for agonist activity with the patched protein. The	
CC	DNA sequence is useful for producing all or portions of the patched	
CC	protein, as probes for research, diagnosis, binding of hedgehog protein	
CC	for its isolation and purification and in gene therapy. The DNA sequence	
CC	is also useful as primers for investigating other species and for	
CC	isolating genes from various mammalian sources of interest, particularly	
CC	from humans or from domestic animals. The DNA sequence is further useful	
CC	for preparing transgenic laboratory animals and to knock out the PTC	
CC	protein in the embryonic stem cells, so as to produce hosts with single	
CC	functional patched gene. The present sequence represents the amino acid	
CC	sequence of the human patched gene PTC product	
XX	Sequence 1447 AA;	
SQ	Query Match 9.4%; Score 651.5; DB 7; Length 1447;	
	Best Local Similarity 22.0%; Pred. No. 2.4e-50;	
	Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;	
QY	225 PQAVGSGIQPLNEGVARCHNESQDDVATCSCQDCAACPAIARPOALDSTFVLGMPGS 284	
DB	24 PGRPAGGRRRTTGGTGLRAAAPDRD-----YLRPSYCDNAFALEQI--- 65	
QY	285 LVLIILICLSVFVAVTILLVGRVAPARDKSKMVDPK-----KGTSLSDKLSFSFTHLL 337	
DB	66 -----SKGATGRKAPLWLRKAFQRLFLFKLCYIQKNC 98	
QY	338 GQFPQGWGTWVASWPLTILVSLVPVVALAAGLVFTLTTDPVELLSAPNSQARSEKAFH 397	
DB	99 GKF-----LVVGLLIFGAFVAGLKAANLETNVEELWVEGVGRVSRRLNYT 143	
QY	398 DOHFGPFFRTN-QVILTAPNRSSRYDSSLGLGPKNFSGILDLDLLELLE--LQERLRHL 454	
DB	144 ROKIGSEAMFNPQMIQTPEEG-----ANVLTTEALLQHLDSALQASRVHV 190	
QY	455 QVMSPEAQRNISLQIDICYAP-----LNPDNTSLYDCCINSLSLQYFQNNRTLLLLTA 505	
DB	191 YMYN-----RQMKLEHLCYKSGELITETGYMDQIIIEVLYPCLIITPLDCFWEGAKLOGTA 246	
QY	506 NOTLMGQ-----TSQVDKDHFLY-----CAN----- 527	
DB	247 --YLLGKPLRLWTNFDPLELELKKINYQVDSWEEMLNKAEVGHGVMRDPCLNPADPDC 304	
QY	528 -----APL-----TPKQGT-----ALALSCMA 544	
DB	305 PATAPNKNSTKPLDMALVNLGGCHGLSRKYMHWQEBELIVGTVKNSGKLVSAHALQTMF 364	
QY	545 DYGAPEPFLAIGGYKGYKDYSEABALIMTFSLNYPAGDPRLOAKOAKLWEAFLEEMRAFQ 604	
DB	365 QLMTPKQMY--EHPKGYEY-----VSHINWNE---DKAAAILLEAWQRTYVEVHVQSV 411	
QY	605 RRMAGMEQVTFPTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSSYSSSRVMVDS 664	
DB	412 AQNSTQKVLSTT--TTLDDILKSFSDVSVIRVAGYLMLAYACLTM---LRWD--CSKS 465	
QY	665 KATLGLGSAVVLGAVMAAMGFFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRL 724	
DB	466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATQVLPFLALGVGVDDVFLLAHAFSET 525	
QY	725 PRPGEPREVHIHGRALGRVAPSMLLCSLSEAI CFFGALTPMPAVRFTALTSLGLAVLDP 784	
DB	526 GONKRIPEEDORTGCLKRTGASVALTSISNVTAFMAALIPALRAFSLQAQVVVVFNF 585	
QY	785 LLQMSAFVALLSLDSKQESRLDVCC-----VKQBEL-----PPPQ 822	
DB	586 AMVLLIPAILSLMDLYRRDRRLDIFCCFTSPCVSRVIOVEPQAYTHTDNTRYSPPPY 645	
QY	823 QGEGL----- 827	
DB	646 SSHFAHETQITMQSTVOLRTEYDPTHVYTTAEPRSEISVQPVTVTQDTLSCQSPST 705	
QY	828 -----LLGFPKAYAPFLLHWITRGVLLFLFLALFGVSL 861	
DB	706 SSTRDLLSQFSDSSLLHCLPECTKWTLLSSFAEKHYAPFLLKPKAKVVVIFLGLGLVSL 765	
QY	862 YSMCHISVGLDQELAPKDSYLLDYFLNRYFEVGAPVYFVTTLGYNFSSEAGMNAICS 921	
DB	766 YGTRVRDGLDLTDIVPRETREYDFIAAQFYFSP-----YNN----- 803	
QY	922 SAGCNMFSTQKIQYAT-----BFPQSYLAIPASS-----WVDDFDWL----- 961	
DB	804 -----YIVTQADYPNIQHLLYDLHRSFNKVKYVMLEENKQLPKMWLHYFRDWLQGLQD 857	
QY	962 -----TPSS-----CCRLYISGPNKDKFCPSTVNSLCLNCKMCSITMGS 1000	
DB	858 AFDSDMETGKMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913	
QY	1001 VRPSVEQPHKYL-PWFLND-----RENIKCPKGLAAY--STSVNLTSDGQ 1043	
DB	914 INPSA--FYIYLTAWVNDPVAYAASQANIRPHRPFWVHDK---ADYMPETRLRIPAAAP 968	
QY	1044 VLASRFMAYHKPLKNSQDYTEALRAARELANITA-DLRKVPCTDPAPEVFPYTTINVPY 1102	
DB	969 IEYAQFPFYLNGLRDTSDFEAIKRVITCSNYTSLGLSSYPNG-----YFP-----LFW 1018	
QY	1103 EGYLILPEGLPMLSLCLVPTFAVSCLLGLDLRSLGNLLISIVMILVDVTGFMALWDIS 1162	
DB	1019 EGYILGRHWLLFISVVLACTFLVCAVFLNFWTAGII-VVYLALMTVELFLGWMGLIGK 1077	
QY	1163 YNAVSLINIVSAGVSVFVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTN 1220	
DB	1078 LSAVPVILIASVGIGVEFTVHALAFLTAIGDKN---RRAVLALEHMPAPVLDG-AVST 1133	

QY 1221 LPIGLVGLAKAQLIOIFPPRLNLLITLGLLGLVPLVSYVG--PDVNPALALEQ- 1277
Db 1134 LLGVLMLAGSEDFIVRYFFAVLAILTILGLVNLGLVLLPVLFFGPFYEVSPANGLNRL 1193
QY 1278 --KRAEEAANVAVSCPNHPSRVSTADNIYVNHSPGSGIKG 1317
Db 1194 PTPSPPPPPSVRFAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 39

ADD46678

ID ADD46678 standard; protein; 1447 AA.

XX AC ADD46678;

XX 02-DEC-2004 (revised)

DT 29-JAN-2004 (first entry)

XX Human Protein NP_000255, SEQ ID NO 12363.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

OS Unidentified.

XX WO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

PR (GEO) GEN HOSPITAL CORP.

PA (PABB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; NP_000255.

XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.

XX Example 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (described in Table 3

CC of the specification) which is differentially expressed during pain.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1447 AA;

Query Match 9.4%; Score 651.5; DB 7; Length 1447;

Best Local Similarity 22.0%; Pred. No. 2.4e-50;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATCSCDCAACPATARPQALSTFFVLGMPGS 284
Db 24 PCRPAAGGRRRTTGLRAAAPDRD-----YLRPSCYCAAFLEQI--- 65

QY 285 LVLIILICSVPAVAVTILLVGPVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKAFORLLFKLGCYQKNC 98

QY 338 GQFFQGWGTWVASWPLTILVSVIPVVALAAGLVFTTELTDPVELWSAPNSQARSEKAFH 337
Db 99 GKF-----LVGGLIFGAFVGLKAANLSTNVEELWVEVGGRVSRRLNT 143

QY 398 DQHFQGFRTN-QVILTAPNRSRYDSLLGPKNPSGILDLDLLELLE--LQELRHL 454
Db 144 RQKIGEEAMFNPQIMIQTPKBERG-----ANVLTTAALQHLDSALQASRVHV 190

QY 455 QVNSPEAQRNISLQDICYAP-----LNPDNTSLYDCCINSLLQYQNNRTLLLLTA 505
Db 191 YMYN-----RQWKLEHLCKYKSGBLITETGYMDQIIBYLFPCLITPLDCFWGAKLQSGTA 246

QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPLRWTFNPDPLEFLELKKINQVDSWEMLNKAEVGHGYMDRCPCLNPADPDC 304

QY 528 -----APL-----TFKQGT-----ALALSCWA 544
Db 305 PATAPNQNSTKPLDMALVILNGCHGLSRKYMHWQBELIVGGTVKNSLGLVSAHALQTMF 364

QY 545 DYGAVPFPPLAIGGVKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEWRAFO 604
Db 365 QLMTPKQMY---EHPKGYEY-----VSHINWNE-----DKAAALEAWQRTVVVVHQSV 411

QY 605 REMAGMFQVTFABRSLEDEINRTTAEADLPFATSYIVIFLYISLALGSYSSWSRVMVDS 664
Db 412 AQNSTQKVLSPFT--TTLDILKSPDSVIRVAGSYLLMLAYACLTM---LRWD--CSKS 465

QY 665 KATLGGGVAVVGLVMAAMGPFYSYLGRSSVLIVQVFPFLVLSVGADNIFIPVLEYQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATTQVLPFLALGVGVDVFLAHAFSET 525

QY 725 PRPGEPREVHIGRALGRVAPSMMLCSLSEALCFELGALTMPMPVARTFALTSLAVILDP 784
Db 526 GQNKRPEDRTEGCEKRTGASVALTSINVTAFPMALIPALRAFSLQAQAAVVVFN 585

QY 785 LIQMSAFVALLSLDSKROEASRLDVCCC-----VKQDEL-----PPPG 822
Db 586 AMVLILFPALLSMDLYREDRLDIFCCFTSPCVSRVIOVEFOAYTDHNTRYSPPPY 645

QY 823 QQEGSL----- 827
Db 646 SSHSPAHEQITMQSTVOLRTYEDPHTVYVYTAEPSEISVQPVTVTQDTLSCQSPST 705

QY 828 -----LLGFPKAYAPFLLHWITRGVLLLLFALFGVSL 861
Db 706 SSTRDLLSQFSDSLHLEPPCTKWTLSFPAEKHYAPLLPKPAKVWIFLGLGLVSL 765

QY 862 YSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTTILGYNFSRAGMAICS 921
Db 766 YGTTVRGLDLDTDVPRETREYDFIAAQFKTFSP-----YNM----- 803

QY 922 SAGCNNFSTQKIQYAT-----BFPESYLAIPASS-----WDDDFIDWL----- 961

Db	804	-----YIVTQAVYENIQHLLYDLHRSFSNVKVMLEENKQLPKMLHYPFRWLQGLQD	857
Qy	962	-----TPSS-----CRLYISGNKDKFCPESTVNSLNCNKMISITWGS	1000
Db	858	AFDSWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913
Qy	1001	VRPSVEQFHXYL--PWFLND-----RPNIKCPKGLLAY--STSVMLTSDGQ	1043
Db	914	INPSA--FVIYLTAMVSDNPVAYAAQANIRPHPEWVHDK---ADYMPETLRIPAAEP	968
Qy	1044	VLASFMAHYKPKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPVEFPPYTTITNVFY	1102
Db	969	IEYAQFPFYLNGLRDTSDFVEAIEKVRTICSNYISLGLSSVPNG-----YFP-----LFW	1018
Qy	1103	EOYLILPEGLPMLSLCLVPTFAVSCLLGLGLDLSRGLNLLSIVMILVDTVGVFMALWDIS	1162
Db	1019	EQYIGLRHWWLLFISVLACTPLVCVAFLLNPWTAGII--VMVLMALTVELFGMGLIGIK	1077
Qy	1163	YNVSLINLVSAVGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1078	LSAVPVILIASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
Qy	1221	LGILVLGLAKAQLIQIFFRNLTLTGLLHGLVELPVLISVYG--POVNPALALBQ-	1277
Db	1134	LLGLVLMAGSEDFVIRYFFAVALITLIGVLNGLVLLPVLISLSPFGPPEVSPANGIARL	1193
Qy	1278	--KRAEEAAVAVMVASCPNHPSRVSTADNIYVNHSPFGSISK	1317
Db	1194	PFPSPFPFVVRFAMPFGHTH--SGSDSSDSEYSSQTTVSG	1233
RESULT 40			
ID	AD94224	standard; protein; 1447 AA.	
XX	AC	AD94224;	
DT	12-FEB-2004	(first entry)	
XX	DE	Human ptc polypeptide.	
XX	KW	Human; patched gene; ptc; developmental abnormality;	
KW	KW	proliferative disorder; tumour; oncogenic patched mutation;	
KW	KW	patched-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma;	
XX	OS	meningioma; fibroma; cancer.	
XX	OS	Homo sapiens.	
XX	FN	US6551782-B1.	
XX	PD	22-APR-2003.	
XX	PF	28-NOV-2000; 2000US-00724631.	
XX	PR	07-OCT-1994; 94US-00319745.	
XX	PR	06-OCT-1995; 95US-00540406.	
XX	PR	31-MAY-1996; 96US-00656055.	
XX	PR	22-AUG-1997; 97US-00918658.	
XX	PA	(STED) UNIV STANFORD.	
XX	PA	(REGC) UNIV CALIFORNIA.	
XX	PI	Scott MP, Goodrich LV, Johnson RL, Epstein E;	
XX	DR	WPI; 2003-615308/58.	
XX	DR	N-PSDB; AD94223.	
XX	PT	Assay for phenotyping patched status of cell in mammalian (preferably	
PT	PT	human) cell sample, involves detecting presence or absence of aberrant	
PT	PT	modification or mutation of a patched gene, and mis-expression of patched	
PT	PT	gene.	
XX	PS	Disclosure; SEQ ID NO 19; 57pp; English.	

XX	CC	The invention relates to an assay for phenotyping the patched status of a		
CC	CC	cell, involving detecting in vitro, in a sample of mammalian cells, the		
CC	CC	presence or absence of a genetic lesion in a patched gene (ptc)		
CC	CC	characterised by at least one of an aberrant modification or mutation of		
CC	CC	a patched gene or mis-expression of the patched gene. The invention also		
CC	CC	relates to a method for diagnosing a proliferative disorder of an animal		
CC	CC	to a developmental abnormality or proliferative disorder correlates with		
CC	CC	developmental abnormality or activity of a patched gene or gene product.		
CC	CC	aberrant expression or activity of a patched gene or gene product.		
CC	CC	involving detecting in vitro the presence of a predisposing mutation in a		
CC	CC	patched gene in cells of the animal. Characterising the phenotype of a		
CC	CC	tumour, involves detecting the presence of an oncogenic patched mutation		
CC	CC	in cells of the tumour, where the presence of the oncogenic mutation		
CC	CC	indicates that the tumour has a patched-associated phenotype. The assay		
CC	CC	is useful for phenotyping the patched status of the cell in a mammalian		
CC	CC	cell sample obtained from a human patient. The disorders that can be		
CC	CC	detected include basal cell nevus syndrome (BCNS), carcinomas,		
CC	CC	meningiomas and fibromas. This sequence represents a human ptc		
CC	CC	polypeptide of the invention.		
XX	XX	Sequence 1447 AA;		
XX	XX	Query Match	9.4%;	Score 651.5; DB 7; Length 1447;
XX	XX	Best Local Similarity	22.0%;	Pred. No. 2.4e-50;
XX	XX	Matches 300; Conservative	181; Mismatches	460; Indels 421; Gaps 47;
Qy	225	PGQAVGSGIQPLNEGVARCHESQGDVATVSCDQCAASCPAIAARPQALDSTFYLGQMPGS	284	
Db	24	PGRPAGGGRRRRTGGLRAAAPDRD-----YIHRPSYCDAAFALEQI----	65	
Qy	285	LVLIIILCSVFVVITLLVGRVAPARDKSKMDPK-----KGTSLSDKLSFSTHTLL	337	
Db	66	-----SKGRATGKAPLMLRAKFORLLFKLGCYIQKNC	98	
Qy	338	GQFFQGWGTWASWPLTILVLSVIPVVALAGLVPTTELTDPVLSWAPNSQARSEKAFH	397	
Db	99	GKF-----LVGLLIFGAFVGLKAAANLETNVEELWVEVWGVRVSRRELYNT	143	
Qy	398	DQHGFPFRFTN-QVLTAPNRSSRYDLSLLGPKNFSGILDLLLELE--LQERLRLH	454	
Db	144	RQKIGEEAMFNPQLMIQPKSEG-----ANVLTEALLQHLDSALQASRVHV	190	
Qy	455	QVWSPEAQRNLSLQICYAP-----LNPNTSLYDCCINSLLQVFNRTLLLLTA	505	
Db	191	YMN---RQMKLEHLCYKSGELITETGYMDQIIEYLPCLITPLDCFWEGAKLQSGTA	246	
Qy	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----	527	
Db	247	--YLLGKPLRWTFNDFLEFLBELKKNYQVDSWEEMLNKAEVGHGYMDRPCINPADPDC	304	
Qy	528	-----APL-----TFKDG-----ALALSCMA	544	
Db	305	PATAPNKNSTKPLDMALVNGCHGLSRKYMWQBELVGGTVKNTGKLVSAHALQTMF	364	
Qy	545	DYCAPVPFPLATGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLEENRAFQ	604	
Db	365	QLMTPKQMY---EHFGYGYEY-----VSHINWNE---DKAAAILLEAWQRTYVGVHQS	411	
Qy	605	RRMAGFMQVTFPAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSSYSSSRVMDVS	664	
Db	412	AQNSTQKVLSTFT--TTLDDILKSFSDSVSVIRVAGSYLLMLAYACLTW---LRWD---CSKS	465	
Qy	665	KATLGLGGVAVVLGAVMAAGFFSYLGRSSVLVQVWPFVLVSAGADNIFIPVLEYSQRL	724	
Db	466	QGAVALGAVLLVALSVAGLGLCSLIGISFNATTOVLFPFALGVGDVDFVLLAHAFSET	525	
Qy	725	PRRPGFREVHIGRALGRVAPSMMLCSLSEACFFPLGALTTPMPAVTPTALTSLAVILDF	784	
Db	526	GQNKRIFFEDRTGECLEKRTGASVALTSISNVTAFMAALIPITPALRAFSLQAAVVVVFNF	585	
Qy	785	LQMSAFVALLSLDSKQEASRLDVCC-----VKQEL-----PPPG	822	

586	AMVLLIPEAILSM	YLRDRRLDIFCCFTSPCVSRVIOVEPOAYTDTHTDNTRYSPPPY	645
823	QGEGL-----		827
646	SSHSAFHEQTQIM	OSTVQLRTEYDPHTVHYVYTTAEPRESEISVQPVTVTQDTLSCQSPST	705
828	-----LLGFFQKAY	APFLHLNITRGVLLFLFALFGVSL	861
706	SSTRDLLSQFSDS	SLHCLPEPCTKWTLSFAEKHYAPFLPKAKVVIIFLFLGLLGVSL	765
862	YSMCHISVLGDOE	LALPKDSVLLDYFLFLNRYFEVGPVYFVTLGVNFSSEAGMAICS	921
766	YGTRVRDGLDIT	VPRETREYDFIAAQPKYFSF-----YNM-----	803
922	SAGCNPFSTKIOY	AT-----EFPEQSLAIPASS-----WDDPFIDWL-----	961
804	-----YIVTKADY	PNIQHLLYDLHRSFSNVKVMLEENKQLPKMHLHYFRDWLQGLQD	857
962	-----TPSS-----	CCRLYISGNKOKFCPCSTVNSLNLKNCXWSITWGS	1000
858	AFDSDWETGKIM	PNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI	913
1001	VRPSVEQHKYL-P	WFLND-----RPNKCPKGGAAV--STSVNLTSDGQ	1043
914	INPSA--FYIYLT	AWVSNPDVAYASAQNIAPRPEWVHDK---ADYMPETRLRIPAAEP	968
1044	VLASRFMAYHKP	LKNSQDYTEALRAARELAANITA-DLRKVPGTDPAPFEVPPYITINVFY	1102
969	IEYAQFPFVLN	GLRDTSDFVEAIEKVRTICSNYTSGLSSYPNG-----YPP-----LFW	1018
1103	EOYLTLPEGLP	MLSLCIVPTFAVSCLLGLDLRSGLNLSIVMLVDTYGFMAWMDIS	1162
1019	EQVIGLRHLLF	ISVLACTELVCVAFLLNPWTAGII-VMLALMTVELFGMGLIGIK	1077
1163	YNVAVSLINLV	SAVGSVEFVSHITRSP--AISTKPTWLERAKEATISMGSVAFGAVAMTN	1220
1078	LSAVPVVILLI	ASVGIGVEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST	1133
1221	LPQILVLGLAK	AQLIQIFFRNLNLITLGLLHGLVFLPVLISYVG--PDVNPALALEQ-	1277
1134	LLGVLMLAGE	PDFVIRYFFAVLAILITLGLVNLGLVLLPVLSPFGPYEVSFANGLNRL	1193
1278	--KRAEAVAAV	MAVSCNHPHRSVSTADNIYVNHSPFGSIKG	1317
1194	PTSPSPPPSV	VRFAFPFGHTH--SGSDSSDSEYSSQTTVSG	1233
RESULT 42			
D	DE48989		
D	AD48989	standard; protein; 1447 AA.	
C	AC	AD48989;	
C	AC	AD48989;	
X	29-JAN-2004	(first entry)	
X	Human patched	protein (ptc).	
X	cytostatic; vul	nerary; gene therapy; phenotyping; patched status;	
X	patched gene; g	enetic predisposition; basal cell nevus syndrome; tumour;	
X	carcinoma; men	ingioma; medulloma; fibroma; cancer; wound healing; aging;	
X	human; patched	gene; ptc.	
OS	Homo sapiens.		
OS	.US2003186309-A1.		
XX	02-OCT-2003.		
XX	22-APR-2003;	2003US-00421446.	
XX	07-OCT-1994;	94US-00319745.	
XX	06-OCT-1995;	95US-00540406.	
XX	31-MAY-1996;	96US-00656055.	

PR	22-AUG-1997;	97US-00918658.		
PR	28-NOV-2000;	2000US-00724631.		
PA	(STRD)	UNIV LELAND STANFORD JUNIOR.		
PA	(REGC)	UNIV CALIFORNIA.		
XX	Scott MP,	Goodrich LV, Johnson RL, Epstein E;		
XX	WPI;	2004-041193/04.		
DR	N-PSDB;	ADE48988.		
XX	Phenotyping the patched status of a cell for diagnosing a genetic			
PT	predisposition for a tumor comprises detecting the presence or absence of			
PT	a genetic lesion having aberrant modification, mutation or mis-expression			
PT	of the patched gene.			
XX	Disclosure; SEQ ID NO 19; 60pp; English.			
XX	The invention describes an assay for phenotyping the patched status of a			
CC	cell comprising detecting in a sample of mammalian cells the presence or			
CC	absence of a genetic lesion having aberrant modification or mutation of a			
CC	patched gene or mis-expression of the patched gene. The assay is useful			
CC	for diagnosing a genetic predisposition of an animal, e.g. basal cell			
CC	nevus syndrome, predisposition for developing tumour, i.e. carcinoma,			
CC	meningioma, medulloma, or fibroma. A genetic construct encoding a patched			
CC	polypeptide is used to treat an animal having a disorder comprising loss			
CC	of function of a wild-type patched gene, such as cancer, and can enhance			
CC	patch function in e.g. wound healing and aging. This sequence is encoded			
CC	by the human patched gene.			
XX	SQ Sequence 1447 AA;			
Query Match				
Best Local Similarity 9.4%; Score 651.5; DB 8; Length 1447;				
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;				
Qy	225	PGQAVSGIOPFNEGVARCNESQGDVATCSQDCAASCAIPARPQALDSTFYLGQMPGS	284	
Db	24	PGRPAGGRRRTGGLRAAARDP-----YLHRPSYCDAAFALEQI----	65	
Qy	285	LVLIILCSVPVAVTILLVGRVAPARDKSMWDPK-----KGTSLSDKLSFSTHTLL	337	
Db	66	-----SKGATGRKAPLMRAKFORLLFKLGCYIQNC	98	
Qy	338	GOFFQGWGTWASWPLTILVLSVIVVALAAGLVFTETDTPVELWSAPNSQASEKAFH	397	
Db	99	GKF-----LVVGLLIFGAFVGLKAAANLETNVEELWVEVGRVSRRELNYT	143	
Qy	398	DOHFGPPFRTH-QVILTAPNRSRYDSLLGPKNFSGILDLILLELE--LQERLRLH	454	
Db	144	RQKIGEEAMFNPQLMIQTPEEG-----ANVLTTEALLQHLDSALQSRVHV	190	
Qy	455	QWVSEAOBNISLQDICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA	505	
Db	191	YMYN-----ROWKLEHCYKSELITETGYMQDIIEYLYPCLITPDCFWEGAKLSGTA	246	
Qy	506	NOTLMGQ-----TSQVDWKDHFY-----CAN-----	527	
Db	247	--YLLGKPLRWTFDPLFLEELKKINQVDSWEEMLNKAEVGHGYMDRPCINPADPDC	304	
Qy	528	-----APL-----TFKDG-----ALALSCMA	544	
Db	305	PATAPNKNSTKPLDMALVNLGCHLSRKYMHQBELVGGTVTKNSTKLVSAHALQTMF	364	
Qy	545	DYGAPVFPPLAIGYKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRAFO	604	
Db	365	QLMTPKQMT---EHFAGYEY-----VSHINWNE---DKAAAILAEAWQTYVEVHVQS	411	
Qy	605	RRMAGMFQVTFPAERSLEDEINRTTAEDLPFATSYVIVIFYLSALGCSYSSNRVWVDS	664	
Db	412	AQNSTQKLSFTT-TLDDILKSFSVSVIRVASGYLLMAYACLTM---LRWD---CSKS	465	
Qy	665	KATLGLGGVAVVLGAVMAAGFFSYLGRSSVILQVWPFLVLSVGADNIFIVLEYQRL	724	

```

Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGVDDVFLLAHAFSET 525
Qy 725 PRPGEPREVHIGRALGRVAPSMMLCSLSEACFFLFGALTPMPAVRTPALTSGLAVILDP 784
Db 526 GQNKRIPEFDRTEGCELRKRTGASVALTSISNVTAFFMAALIPALRAFSLQAVVVVNF 585
Qy 785 LLOMSAFVALLSDSKRQASRLDVCC-----VKQEL-----PPFG 822
Db 586 AMVLFFPAILSLMDLYRREDRLDIFCCFTSPCVSRVIOVEQAYTDTHTNTRYSPPPY 645
Qy 823 QGEGE----- 827
Db 646 SSHSFAHEQITWQSTVQLRTEYDHPHTHYVYTAEPRESISVQPVTVTQDTLSCOSPEST 705
Qy 828 -----LLGFFOKAVAPFLLHMITRQVLLTRGVLLFLALFGVSL 861
Db 706 SSTDRLLSQFSSSLHCLPEPCTKWTLSSFAEKHYAPFLKPKAKVVIFFLGLLGVS 765
Qy 862 YSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGGAPYFVTTLGYNFSSBAGMNAICS 921
Db 766 YGTTRVRDGLDLDIIVPRETREYDFIAAQFKYFSF-----YNN----- 803
Qy 922 SAGCINFSFTKIQYAT-----EPESQSYLAIPASS-----WVDDPIDWL----- 961
Db 804 -----YIVTQADYVNIQHLVDLHRSFSNRYKVMLEENKQLPKMHLHYFRDLQGLQD 857
Qy 962 -----TPSS-----CCRLYISGPNKDKFCPTVNSLNCNKMSITMGS 1000
Db 858 AFSDWETGKINPNYKNGSDGVLAYKLLVQTGSDK--PIDISQLT--KORLVADGI 913
Qy 1001 VRPSVEQFKHYL-PMFLND-----RPNIKCPKGLLAAY--STSVNLTSDGQ 1043
Db 914 INPSA--FYILTAWVNSDPVAVASQAMIRHPPEWHDK--ADYMPETRLRIPAARP 968
Qy 1044 VLASRFMATHKPKNSQDYTEALRAARELANITA-DLRKVGTDPAFEVFPYTTNVFY 1102
Db 969 IBYAQFPFYNGLRDTSDFVEAEKVRTICSNYTSLSGLSSYPNG-----YFP-----LFW 1018
Qy 1103 EGYLTLPGLFWLCLVPTFAVSCLLGLDLRSGLMLLSIVMLVDTVCGFMAWDIS 1162
Db 1019 EGYIGLRHWLLFISVVLACTFLVCVFLNFWTAIGII--VMVLALMTVELFGMGLIGIK 1077
Qy 1163 YNAVSLINLVSAVGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db 1078 LSAVPWVILLASVGIGVEPTVHVALAFLTAIGDKN--RRAVLALEHMFAPVLDG-AVST 1133
Qy 1221 LPGILVLGLAKAQLQIIFFRNLMLITLLGLHLGLVFLPVILSYVG--PDVNPALALBQ- 1277
Db 1134 LLGVLMLAGSEDFIVRYFAVLAITLILGVNLGLVLLPVLSSFFGYPYFEPVSPANGNLRL 1193
Qy 1278 --KRAEEAAVAAVWASCPNHRKSRVSTADNIYVNHSGSIGK 1317
Db 1194 PTPSEPPPPSVVRPAMPQGHTH--SGSDSDSEISQITVSG 1233

```

RESULT 43

AAW52199

ID AAW52199 standard; protein; 1434 AA.

XX

AC AAW52199;

XX

DT 25-JUN-1998 (first entry)

XX

DE Mouse patched (ptc) protein.

XX

KW Patched protein; ptc; cancer; tumour suppressor; cell adhesion promoter;

XX

KW wound healing; ageing; mouse.

XX

OS Mus sp.

XX

FN WO9745541-A2.

XX

PD 04-DEC-1997.

XX

PF 02-JUN-1997; 97WO-US009553.

XX

PR 31-MAY-1996; 96US-00656055.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

PA (REGC) UNIV CALIFORNIA.

XX

PI Scott MP, Goodrich LV, Johnson RL, Epstein E, Oro A;

XX

DR WPI; 1998-032648/03.

XX

DR N-PSDB; AAV21589.

XX

PT Patched protein other than Drosophila melanogaster patched protein - used

PT for characterising the phenotype of a tumour.

XX

PS Claim 5; Page 66-71; 86pp; English.

XX

This is a mouse patched (ptc) protein. The encoding DNA can be used to construct an expression cassette comprising an altered patch or hedgehog gene. The expression cassette comprises a nucleic acid encoding a patched protein other than a Drosophila melanogaster patched protein, or fragment of at least 12 nucleotides in length, as other than an intact chromosome under transcriptional control of a transcriptional initiation region, and a transcriptional termination region, both functional in an expression host. A genetically engineered mammalian cell comprising this expression cassette as an extrachromosomal element or integrated into the genome of the cell can be predisposed to develop basal cell carcinoma as a result of the transfection. By analysing DNA, functional analysis of patched protein function, or by detecting antibody binding to abnormal patched protein, a genetic predisposition to developmental abnormalities and cancer can be diagnosed. This analysis can also be used for characterising the phenotype of a tumour, particularly a carcinoma, especially a basal cell carcinoma. The methods can also be used for characterising transitional cell carcinoma of the bladder, meningiomas medulloblastomas, etc. The modified cells comprising the expression cassette can be used to determine the role of different exons of the patched gene in oncogenesis, signal transduction, etc. Transgenic animal models created from these cells can be used as animal models for carcinomas of the skin. The patched protein of mosquito, butterfly or beetle or alternatively, a mammalian patched protein of human or mouse can be used to identify ligands or substrates that bind to, modulate, or mimic the action of patched gene. These agents could be used as tumour suppressors, cell adhesion promoters (e.g. in wound healing and ageing)

XX

SQ Sequence 1434 AA;

XX

Query Match 9.4%; Score 646.5; DB 2; Length 1434;

Best Local Similarity 22.9%; Pred. No. 7e-50;

Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

Qy 356 LVLVSVIPVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHGFPRFN-QVILTA 414

Db 88 LVWGLLIFGAFVGLKAANLETNVEELVVEVGGVRSRLNVTQKIGSEAMENPQLMIQT 147

Qy 415 PNRSSRYRYSLLGPKNFGSGLDLDLLELLE--LQERLHLQVMSPEAQRNLSQDICY 472

Db 148 PKEEG-----ANVLTTEALLQHLDSALQASRVHYMTN-----ROWKLEHLCY 190

Qy 473 AP-----INPDNTSLYDCCINSLLQYFQNNRTLLLLTANOTLMGO----- 512

Db 191 KSGELITETGYMDQIIIEVYPLCLITPLDCFWEGAKLQSGTA--VLLGKPLRWTFNFDPL 248

Qy 513 -----TSQVDWKOHFLY-----CAN-----APL----- 530

Db 249 EFLEBKIKINQVDSWEEMLANKAERVGHGYMDRCPCLNPADPCPATPNKNSKTKPLDVALV 308

Qy 531 -----TPKQGT-----ALALSCHADYCAPVFPFLAIGYKKG 562

Db 309 LINGCQGLSRKYMHWQEBLIVGTVKNATGKLVSAAHLOTMPQLMTPKQMYHFRGY---- 365

Qy 563 DYSEAEALIMTFSLNNYPAGDPRLAQAQKLWEEAFLEEMRAFORMMAGMFQVTFTAERSLE 622

Db 366 DY-----VSHINNE-----DRAAILEAWQRTVEVHQSVAPNPTQKVLPTT-TTLD 414
Qy 623 DEINRTAEDLPPIATSVIVIFLXISALGSSWSRVMVDSKATLGLGGVAVVLGAVMA 682
Db 415 DILKSFSVSVIRVASGYLLMLAVCLTM-----LWDD--CSKSQAGAVLAGVLLVALSVAA 469
Qy 683 AMGFYSYLGRSSIVLQVQVFLVLSGADNIFIVLEBYQRLPRRPGCEPREVHIGRALGR 742
Db 470 GLGLCSLIGISFNAATTQVLPFLALGVGDVDFLLAHAFSETGQNKRIPEPDRTEGCLKR 529
Qy 743 VAPSMCLSLSEATCFEGALATPMPAVTPTALTGSLAVILDFLLQMSAFVALLSLDSKQ 802
Db 530 TGASVALTSISNVTAFMAALIPALRAFSLQAQVAVVVFAMVLLIFPAILSMDLVRR 589
Qy 803 EASRLDVCCC-----VKQEL-----PPGQGBGL----- 827
Db 590 EDRLDIFCCFTSPCVSRVQVEPQAYTEPHSNTRYSPPPYTHSPFAHETHITMQSTVQ 649
Qy 828 ----- 827
Db 650 LREYDPHTVYVYTTAEPRSEISVQVTVTQDNLSCQSPSTSTRDLSSQFSDSLHCL 709
Qy 828 -----LLGFQKAYAPFLHWTIRGVLLFLALFGVLSYMSCHISVGLDQELALPK 879
Db 710 EPPCTKWTLSSPAERKHYAPFLKPKAKVWVILLFLGLGVSLSYGTTRVRDGLDLTDIVPR 769
Qy 880 DSYLLDYFLFLNRYFEVGPVYVTTTLGVNFSSEAGMNAICSSAGCNFFSFTQKIYAT- 938
Db 770 ETREYDFIAQPKYFSF-----YNM-----YIVTQKADYNI 801
Qy 939 -----EFPEQSYLAIPASS-----WVDDFDWL-----TPSS--- 965
Db 802 QHLLYDLHKSFSNVKVMLEENKQPLQMWLHYFRDWLQGLQDAFSDWETGRIMPENYKN 861
Qy 966 -----CRLYISGNKDKFCPSTVNSLNCNKMSITMGVSRSVEQFHYKL--PWFLN 1017
Db 862 GSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVSN 915
Qy 1018 D-----RPNKCPKGGAAAY--STSVNLTSBGQVLSRFRMAYHKPLKNSQD 1061
Db 916 DPVAYAAQANTRPHRPEWHDK--ADYWPETRLRIPAAEPIEYAAQFPFYNLGLRDTSD 972
Qy 1062 YTEALRAARELANITA-DLRKVPQGTDPAPFVPPYITITNVFEQYITLILPEGLFMLSCL 1120
Db 973 FVEAIEKVRVICNNYTSGLSSYPNG-----YPF-----LFWEQYISLRHWLLSISVVL 1022
Qy 1121 VPTFAVSCLLGLDLRSGLNLLSIWMLVDTVGFNALWDISVNAVSLINLVSAGVMSVE 1180
Db 1023 ACTFLVCAVFLNPWTAGII-VNVLALMTVLEFGMMGLIGIKLSAVPVVILLIASVGIGVE 1081
Qy 1181 FVSHITRSF--AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIF 1238
Db 1082 FTVHVALAFITAGDKN--HRAWLALEHMFAPVLDG-AVSTLLGVLMLAGSEDFIVRY 1137
Qy 1239 FFRNLNLTLLGLLHGLFVLPVLSVG--PDVNPALAEQ-----KRAEBAVAVMWASCP 1293
Db 1138 FPAVALTILTVLGVNLGLVLLPVLSPFGPCPEVSPANGNLRLPTSPPEPPPVVREAVPP 1197
Qy 1294 NHRSRVS-TADNIYVNHSPSGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISELRQYEAQOGAG 1234

RESULT 44

AAW72968

ID AAW72968 standard; protein; 1434 AA.

XX AC

AAW72968;

XX AC

DT 26-JAN-1999 (first entry)

XX AC

DE Mouse patched protein.

XX Mouse; patched gene; diagnosis; treatment; developmental disorder;
KW cancer; healing; injured tissue; ptc; spina bifida; Wnt-1 oncogene;
KW sperm production; gene therapy.
OS Mus sp.
XX US5837538-A.
XX 17-NOV-1998.
XX 06-OCT-1995; 95US-00540406.
XX 07-OCT-1994; 94US-00319745.
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX Scott MP, Goodrich LV, Johnson RL;
PI WPI; 1999-023461/02.
XX N-PSDB; AAV64092.
DR Nucleic acid encoding vertebrate patched protein and related
XX transformants - used to express poly-peptide(s), useful for diagnosis and
XX treatment of developmental disorders or cancer, and in healing of injured
XX tissue.
XX Claim 23; Col 49-56; 38pp; English.
XX The present sequence represents the mouse patched (ptc) protein. Cells
XX containing and expressing the ptc gene are used for the recombinant
XX production of the protein. These in turn are useful: (i) for generating
XX antibodies (Ab); and (ii) to screen for specific-binding ligands
XX (potential therapeutic agonists and antagonists). The ptc gene, or its
XX fragments, are used to isolate related sequences from other mammals; to
XX identify mutations (particularly those associated with genetic disease
XX such as spina bifida and other developmental disorders); to monitor
XX expression levels in testis (to determine relationship with sperm
XX production) and to isolate 5'-non-coding sequences (used to study
XX embryonic development and to provide regulated expression of proteins).
XX The complete gene can be used in gene therapy, including expression of
XX antisense molecules, and to generate transgenic animals for studies of
XX embryonic development. Ab are used diagnostically to determine the ptc
XX protein on cell surfaces and as competitive inhibitors of signal
XX transduction through the ptc ligand. Cells that have been engineered to
XX express the ptc protein can be used to promote regrowth and healing of
XX damaged tissue (e.g. growth of new teeth) and regulation of the ptc
XX protein expression may be useful in cancer treatment (it may control the
XX Wnt-1 oncogene)
XX Sequence 1434 AA;
SQ Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 7e-50;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
Qy 356 LVLSVLPVVALAAGLVFTTELTTDPVLEMSAPNSQARSEKAFHQHFGPPFRTN-QVILFA 414
Db 88 LVVGLLIIFGAFVAGLKAANLETNVELWVEVGRVSRLENTYTRQKIGEAEMFNPLMIQT 147
Qy 415 PNRSSVRYDSSLGLGPKNFSGILDLLLELLE--LQERLRLHQVMSPEAQRNLSQDICY 472
Db 148 PXKEG-----ANVLTEALLQHLDSALQASRVHYVMYN-----RQWKLEHLCY 190
Qy 473 AP-----INPNTSLYDCCINSLLQYFQNNRTLLLLLTANOTLMGO----- 512
Db 191 KSGELITETGYMDQIIIEYLYPCLITTPDCFWEGAKLQSGTA--YLLGPKPLRWTFDPL 248
Qy 513 -----TSQVDWKDHELY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPNLPADPCFATAPNKNSTKPLDVALV 308
Qy 531 -----TFKQGT-----ALALSCMADYAGAPVFPFLAIGYKKG 562

Db 309 LMGCGGLGRKXMHQBEELIVGTGVRNATGKLSAHALQTMFQMTPKQMYEHFRGY--- 365
QY 563 DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEAFLEEMRAFORRMAGMFOVTFPTAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILEAQRYVEVHQSVAPNSTQKVLPTT-TTLD 414
QY 623 DEINRTAEDLPFATSYIVIFLYISALGSYSSSRVMVDSKATGLGGVAVVLGAVMA 682
Db 415 DILKSPFSDSVIRVASGYLMLAYACLTM---LRWD---CSKQCAVGLAGVLVALSVAA 469
QY 683 AMGFPSYLGTRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGPPEVHIGRALGR 742
Db 470 GLGLCSLIGISENAATQTFPLFALGVGDVDFELLAFAFSETQNKRIIPFEDRTGCLXR 529
QY 743 VAPSMELCSLSEAI CFFLGAFTMPAVRTFALTSGLAVIDLDFLOMSAFVALLSLDSKQ 802
Db 530 TGAASVLTISNVTAFPMALPIPALRAFSLQAIVVVFNFAMVLLIPFALLSMDLYRR 589
QY 803 EASRLDVCC-----VKQDEL-----PPQGQEGE----- 827
Db 590 EDRRLDFCFTSPCVSRVIOPEQAYTEPHSNTRYSPPPVYTHSHFAHETHITMQSTVQ 649
QY 828 ----- 827
Db 650 LRTEYDPHTHYVYTTAEPSEISVQPVTVTDNLSCQSPSTSTDRLLSQFSDSLHCL 709
QY 828 -----LLGFQKAYAPFELHWTIRGVVLLFLALFGVSLYSVMCHISVGLDQBLALPK 879
Db 710 EPPCTWTLSSFAEKYAFPLKPKAKVVILLFLGLGVSLYGTTRVRDGLDLDIVPR 769
QY 880 DSVLLDYFLNRYFEVGPVYVFTVLTGYNFSEAGMNAICSSAGCNFSFTQKIOYAT- 938
Db 770 ETRVEDFIAAQKQYFSG-----YNN-----IYVTKADYPI 801
QY 939 -----EFQESYLAIPASS-----WDDFIDWL-----TPSS----- 965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHYFRDMLQGLQDAFSDMETGRIMPNNYKN 861
QY 966 -----CCRLYISGNKDKFCPSTVNSLNCNKCMSITWGSVRPSVEQFHKYL-PWFLN 1017
Db 862 GSDGVLAYKLVQTSRDK--PIDISQLT--KQRLVDADGIIINPSA--FYIYLTAWVSN 915
QY 1018 D-----RPNIKCPKGGLAAY--STSVNLTSQGVQLASRFMAYHKPKNSOD 1061
Db 916 DPVAVAAASQANIRPHEPWHDK---ADYMPETRLRIIPAAEPIEYAFQFPYINGLEDTS 972
QY 1062 YTEALRAARELANITA-DIRKVPQTDPAFVFPYTTITNVFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEAIEKVRVICNNYTSLSGSPNG-----YPF-----LFWEQYISLRHMLLSVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGMALWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCVFLNPTWTAGII-VWVLAJMTVELFGMGLIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTMLERAKATISMSGSAVAGVAMTNLPGILVLGLAKAOLIQIF 1238
Db 1082 FTVHALAFLTAIGDKN---HRAMLAEHMFAPVLDG-AVSTLLGVLMLAGSEFPFIVRY 1137
QY 1239 PFRNLITLGLLHGLVPLVPLSVYG--PDVNPALAEQ---KRAEAAVAVMVASCP 1293
Db 1138 PFAVLAAILTVGLVNLGLVLLPFLVLLSFGFCPEVSPANGLNRLPTSPPEPPSVVRFAVPP 1197
QY 1294 NHPSRVS-TADNINYHNSPEGSI-----KNG 1319
Db 1198 GHTNNGSDSDSEYSSQTIVSGISBELROYEAQQGAG 1234

RESULT 45

AAB67159

ID AAB67159 standard; protein; 1434 AA.

XX

AC AAB67159;

XX 12-APR-2001 (first entry)
XX Murine patched protein.
DE Mouse; patched; PTC; segment polarity; limb patterning; development;
XX hedgehog; antibody; human.
KW Mus sp.
XX US6172200-B1.
PD 09-JAN-2001.
XX 20-OCT-1997; 97US-00954668.
XX 07-OCT-1994; 94US-00319745.
PR 06-OCT-1995; 95US-00540406.
XX (STRD) UNIV LELAND S STANFORD.
XX Scott MP, Goodrich LV, Johnson RL;
XX WPI; 2001-136884/14.
DR N-PSDB; AAF32180.
XX Novel monoclonal antibody useful in diagnostic assays for detection of
PT presence of protein on surface of cells specifically binds to naturally
PT occurring patched protein, other than Drosophila patched protein.
XX Disclosure; Col 47-54; 39pp; English.
XX The present invention provides a monoclonal antibody which specifically
CC binds to a patched protein (PTC) other than that from Drosophila. Also
CC given are the protein and coding sequences of patched from the beetle,
CC mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
CC polarity gene involved in limb patterning. The sequences can be used to
CC study development and to isolate the patched ligand, hedgehog. In
CC addition, antibodies can be used to detect the PTC protein on cell
CC surfaces or to inhibit the transduction of signal by the PTC ligand by
CC competing for its binding site
XX Sequence 1434 AA;
Query Match 9.4%; Score 646.5; DB 4; Length 1434;
Best Local Similarity 22.9%; Pred. No. 78-50;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
QY 356 LVLSVIPVVALAAGLVFTTELTTDPVELMSAPNSQARSEKAFHDQHFPPFRTN-QVILTA 414
Db 88 LVVGLLIFGAPAVGLKAAANLETNVELWVEVGRVSRRELNTYRQKIGBAMFNPQLMIQT 147
QY 415 PNRSYRYVDSLLGPNKPSGILDLDLLELE--LQERLRHLQVMSPEARNISLQDICY 472
Db 148 PKEEG-----ANVLTTEALLQHLDSALQASRVHVVMYN---RQWKLEHLCY 190
QY 473 AP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTANOTLMGQ----- 512
Db 191 KSGELITETGMDQIIIEYLYPCLITITPLDCFWEGAKLOSGETA--YLLGKPPLRWTNFDPL 248
QY 513 -----TSQVMDKDHFLY-----CAN-----APL----- 530
Db 249 EFLBELKXINQVDSWEELNKAEGYCHGYMDRPCLPADPCPATAPKNKSTKPLDVALV 308
QY 531 -----TPKQGT-----ALALSCMADYGAIPVFPFLAIGYKKG 562
Db 309 LMGCGGLSRKXMHQBEELIVGTGVRNATGKLSAHALQTMFQMTPKQMYEHFRGY--- 365
QY 563 DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEAFLEEMRAFORRMAGMFOVTFPTAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILEAQRYVEVHQSVAPNSTQKVLPTT-TTLD 414
QY 623 DEINRTAEDLPFATSYIVIFLYISALGSYSSSRVMVDSKATGLGGVAVVLGAVMA 682

XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX	XX
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

```
QY 563 DYSEBALIMTSLNYPAGDPLRAQAQLWEAPLEEMRAFORMAGMFOVFTTAERSLE 622
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
366 DY-----YSHINWNE-----DRAAILEAQRTYVEVHQSVAPNSTQKVLPTT--TTLD 414
QY 623 DEINRTTADLPIFATSYIVIFLYISLALGSYSSSRVVDKATGLGVAVLGAVNA 682
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
415 DILKSFSDSVIRVASGYLMLAYACLTW---LRMD---CSKQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVLOVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHGRALGR 742
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
470 GLGLCSLIGISENAATTQVLPFLALGVDDVFLLAHAFSETGQNKRIFFEDRTGECLAK 529
QY 743 VAPSMMLCSLSRAICFFICALTPMPAVRTFALTSLGLVLDLQMSAFVALLSDSKRQ 802
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
530 TCGASVALTSISNVTAFPAALPIPALRAFSLQAQAVVVVFNAFVILLIPAILSMDLVRR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPQGQEGE----- 827
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
590 EDRRLDIFCCFTSPCVSRVIOPEQAYTEPHNTRYSPPPPTVSHSFAHETHITMQSTVQ 649
QY 828 ----- 827
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
650 LRTEYDPHTHYVYTTAEPSEISVQPVTVTQDNLSCQSPSTSTSTEDLLSQSDSLHCL 709
QY 828 -----LLGFQKAYAPFLHWTITRGVLLLFALFGVLSYMSCHISVGLDQELALPK 879
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
710 BPECTKWLSSFAEKHYAPFLKPKAKVVVILLFLGLLGLVSLYGTTRVRDGLDLTDIVPR 769
QY 880 DSYLLDYFLFNRYFEVGPVFTVTLGYNFSEAGMNAICSSAGCNPFSTQKIQYAT- 938
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
770 ETREYDFIAAQKYSF-----YNM-----YIVTKADYPMI 801
QY 939 -----EPFEQSYLAIPASS-----WVDDFDIDL-----TPSS--- 965
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
802 QHLLYDLHKSFNKVKYMLEENKQIPOMWLHFRDLQLOQAFDSQDWTETGRIMPNYKN 861
QY 966 -----CRLVISGNKDKPCPSTVNSLNCNKMSITMGSVRPSVQPHKYL--PWFLN 1017
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
862 GSDDGVLAYKLLVQTSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGLAAY--STSVNLSDGQVLSRFAWYHKPLKNSQD 1061
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
916 DPVAAASQANTRPHRPEWHDK--ADYMPETRLRIPAAEPIEYAOQFPFYINGLRDTS 972
QY 1062 YTEALRAARELANITA--DLRKVPCTDPAFEVFPVYITVNFVEQYLTILPEGLFMLSCL 1120
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
973 FVEAIEKVRVICNWTSLGLSYFNG-----YFP-----LFWEQYISLHMLLSISVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFPMALWDISYNVSLINLSVAYGMSVE 1180
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1023 ACTFLVCVAVFLNPTWAGII--VMVLALMTVELFGMGLIGIKLSAVPVVILLIASVGIVE 1081
QY 1181 FVSHITRSF--AISTKPTWLERAKAATISMGSAVAGVAMTMLPGILVLGLAKAQLIQIF 1238
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1082 FTVHVALLAFLTAIGDKN---HRAMLAHEMFAPVLDG-AVSTLLGLVLMAGSEPDFIVRY 1137
QY 1239 PPRMLLITLLGLLHGLFPLVILSVG--PDVNPALALEQ---KRAEEVAAVNVASCP 1293
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1138 FPAVALITVLGLVNLGLVLLPILLSFGCPCEVSPANGLNRLPTSPSPPPSVVRFAVPP 1197
QY 1294 NHPRSVS-TADNIYNHSPGSI-----KGAG 1319
DB || : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1198 GHTNNGSDSDSEYSSQITVSGISELRQYEAQOGAG 1234
```

RESULT 47

ABU62271

ID ABU62271 standard; protein; 1434 AA.

XX AC

ABU62271;

XX DT

25-AUG-2003 (first entry)

XX

DE Mouse patched gene PTC product.

XX Mouse; patched gene; PTC; hedgehog protein; gene therapy.

XX Mus sp.

XX US2003032085-A1.

XX 13-FEB-2003.

XX 20-OCT-1997; 97US-00954701.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX (SCOT/) SCOTT M P.

XX (GOOD/) GOODRICH L V.

XX (JOHN/) JOHNSON R L.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI: 2003-492065/46.

XX N-PSDB; ACA62834.

XX New DNA sequence other than present in a chromosome encoding patched gene other than Drosophila patched gene, useful for preparing transgenic laboratory animals and to knock out patched protein in embryonic stem cells.

XX Disclosure; Page 8-10; 40pp; English.

XX The invention relates to a DNA sequence other than present in chromosome encoding a patched (ptc) gene other than Drosophila patched gene. A cell expressing the DNA sequence is useful for producing patched protein, by growing the cell expressing the DNA sequence, where the patched protein is expressed and isolating the patched protein free of other proteins.

XX The cell expressing the DNA sequence is also useful for screening candidate compounds for binding affinity to the patched protein, by combining the candidate protein with the cell expressing the DNA sequence where the DNA sequence comprises the entire coding sequence under the transcriptional regulation of the transcriptional initiation region and a

XX transcribed region functional in the cell, expressing the patched protein in the cell and assaying for the binding of the candidate compound to the patched protein. The above cell is useful for screening

XX candidate compounds for agonist activity with the patched protein. The DNA sequence is useful for producing all or portions of the patched

XX protein, as probes for research, diagnosis, binding of hedgehog protein for its isolation and purification and in gene therapy. The DNA sequence

XX is also useful as primers for investigating other species and for isolating genes from various mammalian sources of interest, particularly

XX from humans or from domestic animals. The DNA sequence is further useful for preparing transgenic laboratory animals and to knock out the PTC

XX protein in the embryonic stem cells, so as to produce hosts with single functional patched gene. The present sequence represents the amino acid

XX sequence of the mouse patched gene PTC product

XX Sequence 1434 AA;

XX Query Match 9.4%; Score 646.5; DB 7; Length 1434;

XX Best Local Similarity 22.9%; Pred. No. 7e-50;

XX Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVIPVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHGFPRTN-QVILTA 414

DB || :
88 LWGLLTGPAFVGLKAALETNVEELVVRGVRSELTNTROKIGSEANFNPLMIQT 147

QY 415 PNRSSRYDSLGLGPKNFSGILDLDLLELE--LQERLRHQVMSPEAQNISLQDICY 472

DB || :
148 PKEEG-----ANVLTYEALLQHLDSALQASRVHYVYN-----RQWKLEHLCY 190

QY 473 AP-----INPNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMOQ----- 512

DB || :
191 KSGELITGYMDQIIIEVLPCLITPLDCEWEGAKLQSGTA--YLLGKPLPLRTWTFDPL 248

Qy	415	PNRSRYDSSLGKPNFSGILDLJLLELB--LQERLRHLQVMSPEAQRNISLODICY	472
Db	148	PKBEG-----ANVLTTEALLQHLSDALSQASRVHVVMYN---RQWKLEHLCY	190
Qy	473	AP-----LNPDNSTLYDCCNSLLOVQFQNNRTLLLLTANQTLMGQ-----	512
Db	191	KGELIETGYMDQIIEVLYPCLLIITPLDCFWEGAKLQSGTA--YLLGKPLRLWNPDPDL	248
Qy	513	-----TSQVDWKDHFLY-----CAN-----APL-----	530
Db	249	EFLEELKKINYQVDSWEEMLNKAESVGHGYMDRCLNPADPCDPATAPNKNSTKPLDVALV	308
Qy	531	-----TFKQGT-----ALALSCWADYGAPVPFPPLAIGGVKYGK	562
Db	309	LNGCQGLSRKYMWBELIYVGGTVKNAATGKLVSAHALQTMFQMLTPKOMYEHFRGY---	365
Qy	563	DYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLBEWRAPQORMAGMFOVTFPTABRSLE	622
Db	366	DY-----VSHINWNE---DRAAAIIIEAQWRTTVEVHVOSVAPNSTQKVLPFTT--TILD	414
Qy	623	DEINRTAEDLPIFATSYIVIFYLSIALGSYSWSRVMVDKATLGLGGVAVVLGAVMA	682
Db	415	DILKSFSQSVIRVASGYYLLMLAYACLTM---LRWD--CSKSGAGVAGLAVLVALSVAA	469
Qy	683	AMGFPSYLGIRSSLIIVQVVPFLVLSVCAQDNIFIFVLEYQRLPRRCPPEPREVHIGRALGR	742
Db	470	GLGLCSLIGISFNAAATQVLPFLFALGVGVDDVFLLAHAFSETGQNKRIPEDRTEGCLKR	529
Qy	743	VAPSMILCSLEACFFLGALTMPAVRTALTSLGLAVILDFLIOMSAFVALLSLDSKQ	802
Db	530	TGASVALTISINVTAFMAALIPIPALRAFSLOAAVVVFNFAVLLIFPAILSMDLYRR	589
Qy	803	EASRLDVCCC-----VKQDEL-----PPGQGEGL-----	827
Db	590	EDRRLDIFCCFTSPCVSRVIOVEBQAYTEPHSNTRYSPPPPYTSHSPAETHITMQSTVQ	649
Qy	828	-----	827
Db	650	LRTEYDPHTHVYTTABRSISVQPVTVTQDNLSQSPSTSTRDLLSQFSDSSLHLCL	709
Qy	828	-----LLGFFOKAYAPFLHWHITRGVLLLFLALFGVSLYSMSCHISVGLGDELALPK	879
Db	710	EPPCTKWTLSFAEKHYAPFLKPKAKVVILLFLGLLGSVLYGTTVRUGLDLTDIVPR	769
Qy	880	DSYLLDYFLFLNRVFEVGAPVYFTVTLGYNFSSEAGNAICSSAGCNFNFSFTQKIQYAT-	938
Db	770	ETREYDFIAAQFKYFSP-----YNN-----YIVTQKADYPNI	801
Qy	939	-----BPEQSYLAIPASS-----WVDDPFDML-----TPSS---	965
Db	802	QHLLYDLHKFSFNKYVWLEENKOLPQWHLHYFRDLQGLQDAPDSQWETGRIMPNNYKN	861
Qy	966	-----CCRLIYISGPNKDFCPSVTNLSNLCNKMCSITMGSRVPSVEQFHXYL--PWFLN	1017
Db	862	GSDGVLAYKLLVGTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWSVN	915
Qy	1018	D-----RPNIKCPKGGLAAY--STSVNLTSGQVLASRFMAYHPKLNQSQD	1061
Db	916	DPVAYAASQANIRPHRPEWHDK---ADYMPETRLRIPAAEPIEYAFQFPFLNGLRDTSD	972
Qy	1062	YTEALRAARELAANITAT-DLRKVPGTDPAPVRPVTITNNFYEOYLTILPEGLFMLSCLL	1120
Db	973	FVEAIEKRVICNNYTSJGLSSYPNG-----YFF-----LFWEQYISERHWLLLSISVVL	1022
Qy	1121	VPTFAVSCLLGLDLRSLGLNLLSIVMILVDVTGFMALWDISYNAVSLINLVSAGMSVE	1180
Db	1023	ACTFLVCAVFLNPNWTAGI--VMVLAALMTVELFGMWGLIGKLSAVPVVILLIASVGIVGE	1081
Qy	1181	FVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTNLPGIULVGLAKAQILQIF	1238
Db	1082	FTHVVALAFLTAIGDKN---HRAMLALHEHMFAPLDG-AVSTLGLVIMLAGSFPDFTVRY	1137

Qy	1239	PFRNLNLTILGLLHGLVFLPVLISVYG--PDVNPALALRQ---KRAEEAAVAAVMVASCP	1239
Db	1138	FFAVLAITLVGLVGLVLLPVLISFPGCPCEVSPANGNLRLPTSPPEPPSVVRFAVPP	1197
Qy	1294	NHPSRV--TADNLYVNHSPFGSI-----KGAG	1319
Db	1198	GHTNNGSDSDSEYSSQTVTSGISELRQYEAQQGAG	1234
RESULT 49			
ADH62722			
ID	ADH62722	standard; protein; 1434 AA.	
XX	ADH62722;		
XX	AC		
XX	25-MAR-2004	(first entry)	
DT	Mouse	patched (mptc) protein.	
DE	Mouse	patched	
XX	PTC; patched;	embryonic development; cellular regulation;	
KW	signal transduction;	mouse.	
KW			
XX	Mus	sp.	
OS			
XX	US2003148388-A1.		
FN			
XX	07-AUG-2003.		
PD			
XX	03-JAN-2001;	2001US-00754032.	
PF			
XX	07-OCT-1994;	94US-00319745.	
PR	06-OCT-1995;	95US-00540406.	
PR	20-OCT-1997;	97US-00954668.	
XX	(STRD)	UNIV STANFORD.	
PA			
XX	Scott MP,	Goodrich LV, Johnson RL;	
PI			
XX	WPI; 2003-897566/82.		
DR	N-PSDB; ADH62721.		
DR			
XX	New DNA sequences encoding patched protein from organisms including human		
PT	and mouse are used to elucidate embryonic development and cellular		
PT	regulation associated with patched signal transduction and identify		
PT	patched ligands.		
XX	Disclosure; SEQ ID NO 10; 40pp; English.		
ES			
XX	The present invention relates to a DNA sequence encoding a patched (PTC)		
CC	gene. The invention is used to elucidate embryonic development, cellular		
CC	regulation associated with signal transduction by the patched gene,		
CC	identify agonists and antagonists of that signal transduction, identify		
CC	and isolate ligands of patched protein and assay for levels of		
CC	transcription of the patched gene. The present sequence is the mouse		
CC	patched protein.		
XX			
XX	Sequence 1434 AA;		
Query Match	9.4%;	Score 646.5; DB 7; Length 1434;	
Best Local Similarity	22.9%;	Pred. No. 7e-50;	
Matches 283; Conservative	171; Mismatches 420; Indels 363; Gaps 44;		
Qy	356	LVLSPVPPVVALAAGLVFELTTPDVELWSAPNSQARSEKAFDHFQGFPRFN-QVILTA	414
Db	88	LVGGLLIFGAPVGLKKAANLTVNELWVEVGRVSRRELNTYTKIGEEAMFNPQLMIQT	147
Qy	415	PNRSSVRYDSILLGPKNFSGILDLDLLELLE--LQERLRHLQVWSPQAQRNISLQDICY	472
Db	148	PKEEG-----ANVLITEALQHLDSALQASRVHVVMYN----RQWKLEHLCY	190
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMGQ-----	512
Db	191	KSGELITETGVMDOIIEYLPCLIIITPLDCFWEGAKLOSCTA--YLAKGKPLRWTFNDFPL	248

Db 191 KSGELITETGYMDQIIIEYLYPCLIITPLDLCFWEGAKLQSGTA--YLLGKPPPLRWTFNFDPL 248
Qy 513 -----TSQVMDKDFLY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPNADPCDPATAPNKNSTKPLDVALV 308
Qy 531 -----TPKDG-----ALALSCMADYGAPVPPFLAIGYKKG 562
Db 309 LNGGCGGLSRKYMHWQBELIVGGTVKNATGKLSAHLQTMFLQMTPKQMYHFRGY--- 365
Qy 563 DYSEAEALITWFSLNYPAGDPRLAQAKLBEAFLEEMRAFORRMAGMEQVTFTEARSLE 622
Db 366 DY-----VSHINWE-----DRAAILAEAWQTYVEVHQSVAPNSTQKVLPEFTT--TLLD 414
Qy 623 DRINKRTABDLPIFATSYIVIFLYISLALGSYSMSRVMDSKATILGLGGVAVLIGAVNA 682
Db 415 DILKSFSDVSVIRVAGVLLMLAYACLTW--LRWD--CSKQGAVALGLVLLVALSVAA 469
Qy 683 ANGPFYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHIGRALGR 742
Db 470 GLGLCSLIGISFNAATTQVLPFALGVGVDDVFLLAHAFSETGQNKRIIPFEDRTGECLNR 529
Qy 743 VAPSMMLCSLSEAIQFELGALTPMPAVRTFALTSLGLAVILDLLQMSAFVALLSDSKRQ 802
Db 530 TGASVALTSISNWTAFMAALIPIPALRAFSLOAAVVVFNAMVILLIPFALISMDLYNR 589
Qy 803 EASRLDVCC-----VKPQEL-----PPGQEGEL----- 827
Db 590 EDERLDIFCFTSPCVSRVIOEQPAYTEPHSNTRYSPPPYTSFPAETHITMQSTVQ 649
Qy 828 ----- 827
Db 650 LRTEYDPHTHYTTAEPREISVQPVTVTDNLSCQSPSTSTRDLLSQFSDSLHCL 709
Qy 828 -----LLGFQKAYAPFLHWITRGVILLFLALFGVSLYSMSCHSVGLDQELAPK 879
Db 710 EPCPKTWLSSFAERHAYPFLKPKAKVGVILLGLLGVSLYGTTRVRDGLDLTDIVR 769
Qy 880 DSYLLDYFLNRYFEVGPAPVFTVTLGYNFSSSEAGMNAICSSAGCNNSFTQKIYAT- 938
Db 770 ETRVDYFIAAQKYSP-----YMW-----YIVTKADYPI 801
Qy 939 -----EPPEQSYLAIPASS-----WDDFDWL-----TPSS--- 965
Db 802 QHLLYDLHLKFSNVKYMLEENKQLPQMWLHYFRDWLQGLQADFSDWETGRIMPNYKN 861
Qy 966 -----CCRLYISGNKDKFCPSTVNSLNCIKNCHSTMGSVRPSVEQPHKYL--PWFLN 1017
Db 862 GSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIIINPSA--FYIYLTAWVSN 915
Qy 1018 D-----RPNIKCPKGLAAY--STSVNLTSDDGOVLASRFMAYHKPKNSQD 1061
Db 916 DPVAVAAQANTRPIRPEWHDK---ADYMPETRIPIAEPIEYAPQPPYNGLRDUSD 972
Qy 1062 YTEALRAARELANITA--DLRKVPGTDPAPFVFPYTYITNVFYEQYLLTILPEGLFMLSCL 1120
Db 973 FVEAIEKRVICNNVYTSGLSSYPNG-----YPP-----LFWEQYISLHMLLSISVVL 1022
Qy 1121 VTFPAVSCLLGLDLRSGLNLLSIVMILVDTVGFMALMDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCAVFLNPNWTAGII--VMVLAMTVFELFGMGLIGIKLSAPVPPVILLIASVGIGVE 1081
Qy 1181 FVSHLTRSF--AISTKPTWLERAKATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIF 1238
Db 1082 FTVHVALAPLTAIGDN-----HRAMLALEHMFAPVLDG--AVSTLLGLVLMLAGSEFFIVRY 1137
Qy 1239 FFRLLMLITLLGLLHGLVPLVILSYVG--PDVNPALALEQ---KRAEAAVAAVMVASCP 1293
Db 1138 FPAVALITVLVGLNGLVLLPVLLSFFGFCPEVSPANGLNRLTPSPPEPPSVVPAVPP 1197
Qy 1294 NHPRSVS--TADNIYNVHNSPEGSI-----KGAG 1319
Db 1198 GHTNNGSDSSDSESSQTTVSGISBELRYEAOQAG 1234

RESULT 51

AAR94380

ID AAR94380 standard; protein; 1434 AA.

XX

AC AAR94380;

XX

DT 11-JUL-1996 (first entry)

XX

DE Mouse patched protein.

XX

KW Patched gene; PTC protein; embryo development; cellular regulation; signal transduction; ligand; antibody; hedgehog protein.

XX

OS Mus musculus.

XX

PN WO9611260-A1.

XX

PD 18-APR-1996.

XX

PP 06-OCT-1995; 95WO-US013233.

XX

PR 07-OCT-1994; 94US-00319745.

XX

PA (STRD) UNIV LELAND STANFORD JUNIOR.

XX

PI Scott MP, Goodrich LV, Johnson RL;

XX

DR WPI; 1996-209842/21.

DR

N-PSDB; AAT14218.

XX

PT DNA encoding patched protein other than Drosophila patched protein - used to produce antibodies which detect or inhibit patched protein ligand

PT

XX signal transduction in cells.

XX

PS Disclosure; Page 46-51; 70pp; English.

XX

CC Mouse patched protein (PTC) (AAR94380) was identified as the product of a cDNA clone (AAT14218) derived from mouse limb bud. It has about 38% identical amino acids to Drosophila PTC. PTC protein has been proposed as a receptor for hedgehog protein on basis of genetic experiments in flies. Murine PTC protein can be obtd. in large amounts by expression of the cDNA clone in transformed host cells. It can be used to screen for agonists and antagonists, to isolate its ligand, partic. Sonic hedgehog, to assay for the transcription of ptc mRNA and to raise antibodies

XX

SQ Sequence 1434 AA;

Query Match

Best Local Similarity 9.3%; Score 643.5; DB 2; Length 1434;

Matches 282; Conservative 172; Mismatches 420; Indels 363; Gaps 44;

Qy 356 LVLSVIPVVALAAGLVFTTLPVLSAPNSQARSEKAFHQHFGPFFRTN-QVILPA 414

Db 88 LVVGLLIIFGAFVGLKAAANLETNVEELWVEVGRVSRRLNVTROKIGSEAMFNPQLMIQT 147

Qy 415 PNRSSRYDSLLGLPNFSGILDLLLELLE--LQERLRLHLOWSPQAQRNLSLODICY 472

Db 148 PKBEG-----ANVLTTBALLQLHLSALQASRVHYMYN----RQWKLEHLCY 190

Qy 473 AP-----LNPDTSLYDCINSLLYQFNQNTLLLLTANQTLMGQ----- 512

Db 191 KSGELITETGYMDQIIIEYLYPCLIITPLDLCFWEGAKLQSGTA--YLLGKPPPLRWTFNFDPL 248

Qy 513 -----TSQVMDKDFLY-----CAN-----APL----- 530

Db 249 EFLEELKKINYQVDSWEEMLNKAEVGHGYMDRPNADPCDPATAPNKNSTKPLDVALV 308

Qy 531 -----TFKDG-----ALALSCMADYGAPVPPFLAIGYKKG 562

Db 309 LNGGCGGLSRKYMHWQBELIVGGTVKNATGKLSAHLQTMFLQMTPKQMYHFRGY--- 365

QY 478 DNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQ--VDWKDHFYLCANAPLTFPGGT 536
Db 152 D-----CPATAPNKNSTKPLDMALVTLNGCGCHLSRKYMHMOBELIVGQ-----TVKNST 200
QY 537 -----ALALSCMADYCAPVPFPLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKL 591
Db 201 GKLVSAHALQTMPLTPQMY-----EHPKGYEY-----VSHINWNE-----DKAAALISA 247
QY 592 WEBAFLEEMRAFORMMAGMFQVFTTAERSLEDSINRTTAEDLPFATSYVIFLYISLAL 651
Db 248 WQTYVEVWHQSVAQNSTOKVLSFTT--TTLDLILKXSFSDSVIRVASGYLLMLAYACLTM 306
QY 652 GSYSSRSRVWDSKATLGGVAVLVGAVMAAGPFYSYLGIRSLVILQVVPFLVLSVGA 711
Db 307 ---LRWD---CSKSGQAGVGLVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGV 361
QY 712 DNIFIEVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEAI CFFELGALTPEMAVPT 771
Db 362 DDVFLLAHAFSETGQNKRIPFEDRTGECLEKRTGASVALTISINVTAFMAALIPALRA 421
QY 772 FALTGLAVILDFLLQMSAFVALLSDSKRQBSRLDVCCC-----VKQDEL-- 818
Db 422 FSLQAAVVVVVFAMVLLIFPALLSMDLYRREDRLDIFCCFTSPCVSRVIQVEPQAYTD 481
QY 819 -----PPQCGBEL----- 827
Db 482 THDNTRYSPPPYSSHSFAHETQITMQSTVQLRTEYDPHTHVYVYTTAEPREBSIQPVTV 541
QY 828 -----LLGFFOKAYAPFLLHMTTRGV 848
Db 542 TODTUSCQPESTSTRDLSLQSFSSSLHCLBPCTKWTLSFSAEKHYAPFLLPKPKVV 601
QY 849 VLLFLFALFGVSLYCHISVGLDQELAPKDSYLLDYFLNRYFVEGAPVYVFTLGY 908
Db 602 VIFLFLGLGVSILYGTTRVDGLDLTDIVPRETRVDFTAAQPKYFSP-----Y 650
QY 909 NFSSEAGMNAICSSAGCNFPSTQKIQYAT-----EPPEOSYLAIPASS-----W 953
Db 651 NM-----YIVTQADYPNIQHLLYDLHRSFSNVKYVLMLENKQLPKMW 693
QY 954 VDDFIDWL-----TPSS-----CCRLYISQPNKDKCPTVNSL 987
Db 694 LHYFRDMLQGLQADPDSWETGKIMPNYKNGSDGVLAYKULLVOTGSRDK--PIDISOL 751
QY 988 NCLXNCMSITGSRVPSVQFHKYL--PWFLND-----RPNIKCPKGLAAY 1032
Db 752 T--KQLVDADGIINPSA--FYILTAWVSNPDVAYASQANIRPHRPWVHDK--ADY 804
QY 1033 ---STVNLTSDGOVLASRFMAHYKPKNSQDYTEALRAARELAANITA--DLRKVPGTDA 1089
Db 805 MPETRLRIPAAEPFIEYAQPPFYVLNGLRDTSDFVEAIEKVRTICSNVTSLSGLSSYPNG--- 861
QY 1090 FEVPFPTITNVFYEQVLTILPGLFMLSCLVPTFAVSCLLILGLDRSLNLLSIVMLL 1149
Db 862 ---YFP---LFEWQYIGRHWLLLFISVVLACTFLVCAVFLNPNWTAGII--VMVALMT 913
QY 1150 VDTVGFMALWDISYNAVSLINLVSAVMSVFVSHITRSF--AISTKPTWLRRAKEATIS 1207
Db 914 VELFGMMGLIGIKLSAVPVVILLIASVGIVETVHVALAFLTAISDKN---RRAVLALEH 970
QY 1208 MGSAPVAGVAMTNLPGILVLGLAKAQLIOIFPFRNLNLTLLGLHLGLVFLPVILSYVG- 1266
Db 971 MPAPVLDG-AVSTLLGLVLMAGSEPDFIVRYEFAVLAAILTILGVNLGLVLLPVLSPFPG 1029
QY 1267 -PDVNPALALEQ---KRAEEAANAAMVASCNHPRSRVSTADNIYVNHSPGSIKG 1317
Db 1030 YPEVSPANGNLNLTPTSPPEPSPVVRFAAMPGGHTH--SGSDSSDSSEYSSQTTVSG 1082

RESULT 53

ADI21144

ID ADI21144 standard; protein; 967 AA.

XX

AC ADI21144;

XX 15-APR-2004 (first entry)

XX Novel human protein #119.

XX forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.

OS Homo sapiens.

XX WO2003025148-A2.

XX 27-MAR-2003.

XX 19-SEP-2002; 2002WO-US0299964.

XX 19-SEP-2001; 2001US-0323739P.

PR 13-SEP-2002; 2002US-00323739.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;

PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;

PI Haley-Vicente D;

XX WPI; 2003-354603/33.

DR N-PSDB; ADI21860.

XX New polynucleotides and secreted proteins, useful for treating myeloid or
PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
PT tissue growth or regeneration, in wound healing, and in tissue repair and
PT replacement.

XX Claim 20; SEQ ID NO 395; 156pp; English.

XX The invention relates to an isolated polynucleotide encoding a
CC polypeptide with biological activity. The polynucleotides and
CC polypeptides are useful in diagnostics, forensics, gene mapping,
CC identification of mutations responsible for genetic disorders and other
CC traits, to assess biodiversity, as nutritional sources or supplements.
CC The polynucleotides may also be used as molecular weight markers.
CC chromosome markers or map related gene positions, or as an antigen to
CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
XX protein.

XX Sequence 967 AA;

Query Match 8.5%; Score 588.5; DB 7; Length 967;

Best Local Similarity 21.3%; Pred. No. 9.8e-45;

Matches 216; Conservative 192; Mismatches 359; Indels 249; Gaps 37;

QY 334 HT-----LLGQFFQCGMGWTVASWPLTILVLSVIPVVAALAAGVF--TELTDPVLSWAP 386

Db 115 HTDCLGLSRTFQWLGMQVGAHPWIFLAPLMLTAALGTGLYLPKDEEDLESHYTPV 174

QY 387 NSQARSEKAFHDQHFPGPFRTNQVILTAPNRSS-----YRYSLLLPKPNFSGIL 436

Db 175 GSPAKAERRFQGH-----FTTNDYRFSASRSTANFVSLVVSYSLSLLDPATFAEVS 230

QY 437 DLDLLELLELLEQLERLHQLQWSPEAQRNLSIQDIC--YAPLNPDNTSLYDCCI--NSLLQ 492

Db 231 KLDGAVQDLRV-----AREKGSQIQQVCARVRAI-----CVPPNPILY 270

[illegible]

cytostatic; testicular cancer; hedgehog protein signalling.

Homo sapiens.

Key Location/Qualifiers
Domain 57..76 /label= TM1
Domain 394..414 /note= "Transmembrane domain"
Domain 427..449 /label= TM2
Domain 458..477 /note= "Transmembrane domain"
Domain 502..524 /label= TM3
Domain 532..553 /note= "Transmembrane domain"
Domain 567..705 /label= TM4
Domain 594..984 /note= "Transmembrane domain"
Domain 988..1010 /label= TM5
Domain 1014..1033 /note= "Transmembrane domain"
Domain 1058..1077 /label= TM6
Domain 1091..1114 /note= "Transmembrane domain"
Domain 1091..1114 /label= TM7
Domain 1091..1114 /note= "Transmembrane domain"

US6348575-B1.

19-FEB-2002.

15-APR-1999; 99US-00293505.

15-APR-1998; 98US-0081884P.

(GETH) GENENTECH INC.

De Sauvage P, Carpenter DA;

WPI; 2002-215260/27.

N-PSDB; AAD31576.

Native human patched-2 polypeptide for treating disorders caused by Hedgehog protein signalling such as testicular cancer, and for screening cDNA libraries.

Example 1; Fig 1; 82pp; English.

The invention relates to an isolated sequence comprising a native human patched-2 (Ptc-2) polypeptide. The invention also relates to signalling and mediator molecules in the hedgehog (hh) cascade which are involved in cell proliferation and differentiation. The isolated sequence is useful for the treatment of disorders which are linked to Hedgehog, especially Desert hedgehog expression, such as testicular cancer. It may also be used as a hybridisation probe in a cDNA library to isolate Ptc-2 or its homologues, and to diagnose whether a disorder is driven by Ptc-2 or Hedgehog protein signalling. The present sequence is human patched-2 (Ptc-2) protein

XX	Sequence 1203 AA;	Query Match
SQ		Best Local Similarity 22.9%; Pred. No. 3.8e-44; Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
QY	306 RVAPARDKSKMVDPKKGTSLSDKL--SFSTHTLLGQFFQG--WGTWVASPLTTLVLVS	8.5%; Score 584; DB 5; Length 1203;
DB	3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLPSLGGCQIRHCGKVLFLG	
QY	360 VIPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHDQHF--PFFPTNVOILTAPNR	
DB	63 LLAFGALALGLRMAIETNLEQLWVEGSRVSRVQELHYTKELGEEAAVTSQMLIQTAQRE	
QY	418 SSVRYDSLGLGPKFNSGILDLDLLELELEQLRHLQVMSPEAQRTNISLQIDICY--AP	
DB	123 GEN-----ILTPALG-----LHLQAULTASKVQVSLYKGSMDLNKICYKSGVP	
QY	475 LMPDN-----TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKD-----	
DB	167 LIENGMIEMIEKLPCCVILITPLDCFWEGAK---LQGSAYILPGRPDIOQTNLDPEQLLE	
QY	521 -----HFLYCANAPLTFKDGFTALALSCMADYGA	
DB	224 ELGPPASLEGFRELDDKAQVQVYVGRPCLHPDDLHCPPSAPNHHSRQAPNVAAHLSGGC	
QY	549 PVFPF-----LAIGGVKGYSE--ARALIMTF-----SLNYPAGDPRLAQAKL	
DB	284 HGFSHKFMHQBELLGGWARDPQGBELLRAELAQSTFLMSPRQYEHPRGDTYTHDIGW	
QY	592 WEEAFLEEMRAFQRMAGMFQVTFATERSLE--DEINRTTAEDLPFATS-----Y	
DB	344 SEEQASTVLQAWQRFVQLAQEALPENASQIHAFTSTLDDI--LHAISEVSAARVVGYY	
QY	641 IVIFLYISLALGSYSSWSRVMDSKATIGLGVAIVLGAVMAMGFFSYLGRSSLVILQ	
DB	403 LLMLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGICALLGITFNAAATQ	
QY	701 VVPELVTSVGADNIFIVLEYQRLPRRGRPREVHIGRALGRVAPSMLLCSLSEACFPL	
DB	458 VLPFLALGIGVDDVFL--LAHAFTALPGTFLQERMGECLQRTGTSTSVLTSINNMAAFIL	
QY	761 GALTMPAVRTFALTSLGLAVILDFLLOMSAFVALISLDSKQESRLDVCCC-----	
DB	516 AALVPIPALRAFSLOAAIVVGCTFVAVMLVPPAILSLDLRRHRCQLDLVCCFSFCSAQ	
QY	813 ---VKPQEL-----PPPGQSEGL--	
DB	576 VIQLPQELGDTVPVGIAHLTATVQAFTHCEASSQHVVTILPPQAHLPVPPSPDLGSEL	
QY	828 -----LLG-----PFQKAYAPFLHWTIRGVLLPLA	
DB	636 FSPGGSTRDLILQGBEETROKAAKSLPCARNLAHFARYQFAPLILQSHAKAIVLVLFGA	
QY	856 LFGVSLYSMCHISVGLDQELALPKDSYLLDFLNRVFEVGAPVYFVTTLGYNFS--SEA	
DB	696 LLGLSLYCATLVQDGLATDVVPGTKEHAFSLAQRLYFSL--YEVALVTQGGFDYAHSGR	
QY	915 GMAITCSSAGCNFSTQKIQYATEFPBQSYLAIPAS-----SWVDDFDIMLTPSSCC--	
DB	755 AL-----FDLHQRFSSL-----KAVLPPTATQAPRTWLHYRNWLQGIQAAPDQ	
QY	968 -----RLYISGPNKDKFCPTVNSLNCNKCMSTWGSVRPSVQFPHKYLFWFLND	
DB	799 DWASGRITRHSYRNG--SED-----GALAYKLLIQTDGAQEPDL--FSQLATTKLVD	
QY	1019 RPNIKCPKGGGLAAVSTSVNLTSDGOVLASRFMAVHKP-----	
DB	847 REGLIPEE--LFWYMLTIVWSSDPLGLAASQANFYPPPPPEWLHKYDTTGEMLRIPPAQ	
QY	1056 -----LKNSQDYTBALRAARELANI--TADLRKVPGTDPAPFVFPYTTITNVFY	

Db 905 LEFAQFPPLRLGLOKTADEFVAIEGARACABACQAGVHAYPSGSPF-----LFW 954
Qy 1103 EQYLTPLEGFLMISLC--LVPTEFAVSCLLGLDLRGLMLLSIVMLVDTVGFMALWD 1160
Db 955 EQYGLG--RRCFLLAVCLLVCTFLVCALLLNPNWAGLI-VLVLAWMVTELFINGFLG 1011
Qy 1161 ISYNAVSLINVASGMSVEFVSHITRSFALSTKPTWLERAKEATISNGSAVFAGVAMTN 1220
Db 1012 IKLSAIPVILVASGIGVEFTVHALGF-LTTQGSRLNRAHALEHTFAPVTDG-AIST 1069
Qy 1221 LPGILVLGAKAQLQIFERLNLITLLGLLHGLVFLPVLISVVGPDVNPALALEOKRA 1280
Db 1070 LGLLMLAGSHPDFIVRFFAALTVTLGLLHGLVLLPVLISILGPP-PEVIQMYKES 1127
Qy 1281 BEAANA-----MVASCPNHPRSRVSTADNIYN 1308
Db 1128 PILSPAPQGGGLRWGASSLPQSFARVTTSMTVAIH 1165
XX ABG74104 standard; protein; 1203 AA.
XX ABG74104;
XX 02-APR-2003 (first entry)
XX Human patched-2 protein.
XX Human; patched-2; Dhh signalling; proliferation; differentiation;
XX Chromosome 1p33-34; male contraceptive; infertility; spermatogenesis;
XX basal cell carcinoma; neurodegenerative disorder; memory deficit;
XX Alzheimer's disease; Parkinson's disease; Lou Gehrig's disease;
XX Huntington's disease; schizophrenia; stroke; drug addiction; gut disease;
XX bone disease; skin disease; testicular disease; ulcer; lung disease;
XX pancreatic disease; diabetes; osteoporosis; desert hedgehog.
XX Homo sapiens.
XX OS
XX US2002156245-A1.
XX 24-OCT-2002.
XX 20-NOV-2001; 2001US-00990046.
XX 15-APR-1998; 98US-0081884P.
XX 15-APR-1999; 99US-00293505.
XX (GETH) GENENTECH INC.
XX De Sauvage FJ, Carpenter DA;
XX WPI; 2003-182650/18.
XX N-PSDB; ABX15919.
XX New nucleic acid, useful for manufacturing a medicament for diagnosing or
XX treating a disorder that is modulated by Desert hedgehog (Dhh) signaling
XX e.g., testicular cancer.
XX Claim 14; Fig 1; 85pp; English.
XX The invention relates to a new isolated nucleic acid encoding a
XX polypeptide having patched-2 biological activity, comprises DNA having at
XX least 95% sequence identity with a DNA molecule or its complement
XX encoding: (a) a human patched-2 polypeptide comprising the sequence
XX 'ABC74104'; or (b) the same mature polypeptide encoded by the cDNA (ATCC
XX Deposit No. 209778 designation). Also included are a vector comprising
XX the nucleic acid, a host cell transformed with the vector, a process for
XX producing patched-2 polypeptides, an isolated native sequence of human
XX patched-2 polypeptide, a chimeric molecule comprising the vertebrate
XX patched-2 polypeptide and a heterologous (aa) sequence, an antagonist of
XX patched-2 (that blocks, prevents, inhibits and/or neutralises the Desert
XX hedgehog (Dhh) function in the Dhh signalling pathway), an agonist of

CC patched-2 that stimulates or enhances the normal functioning of patched-2
CC in the Dhh signalling pathway, screening for ant/agonists of patched-2
CC and diagnosing to determine whether a particular disorder is modulated
CC The nucleic acid is useful for manufacturing a medicament for diagnosing
CC or treating a disorder that is modulated by Dhh signalling e.g. tumour,
CC basal cell carcinoma, neurodegenerative disorders, memory deficit,
CC Alzheimer's disease, Parkinson's disease, stroke, drug addiction, gut
CC Huntington's disease, schizophrenia, diabetes, testicular diseases, lung
CC diseases, bone diseases, skin diseases, osteoporosis and infertility.
CC Patched-2 may also be used to develop male contraceptives. The gene for
CC patched-2 is located on human chromosome 1p33-34. The present sequence
CC represents human patched-2 protein
XX
XX Qy Sequence 1203 AA;
Qy Query Match 8.5%; Score 584; DB 6; Length 1203;
Db Best Local Similarity 22.9%; Pred. No. 3.8e-44;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
Qy 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGQFQG---WGTWVASWPTLILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQSLGLGCIQRHCKVFLFG 62
Qy 360 VIPVVALAAGLVFTLTDTPVELWSPNSQARSEKAFHDQHFQ--PFRTNQVILTPNR 417
Db 63 LLAFGALALGLRMAIETNLEQLWVEGSRVSOELHYTKELGEEAAYTSQMLTQAROE 122
Qy 418 SSRYDLSLLGPNFSGILDLLELLELQERLHLQVMSPEAQRNLSQIDICY---AP 474
Db 123 GEN-----ILTPEALG-----LHQAALTASKVQSVLSYKSWDLNKKICYKSGVP 166
Qy 475 LNPDN-----TSLYDCCINSLLQYFQNNRTLLLLLTANTQTLMGQTSQVDKX----- 520
Db 167 LIENGMEIWMIEKLPFPCVILLTDLPCFWEKAK---LQGSAYLPGRPDIQNTLNDPEQLLE 223
Qy 521 -----HFLYCANAPLTFKDGDTALALSCWADYGA 548
Db 224 ELGPPASLEGPRELLDKAQVQAVYGRPCLHPDLDLHCPSPAPNHHRSQAPNAHELSGGC 283
Qy 549 PVFPF-----LAIGYKGDYSE---AEALWTP-----SLNNYPAGDPRLAQAKL 591
Db 284 HGFSHKFMHQBELLGGMARDPQCELLRALALQSTFLLMSPRLQYEHFRGDYQTHDIGW 343
Qy 592 WEEAFLEEMRAFORRMAGMFQVTFTAERSLE--DEINRTTABEDLPFAFS-----Y 640
Db 344 SEEQASTVLQAWQRFFVQLAQEALPENASQIHAFSSITLDDI-LHAFSEVSAARVWGY 402
Qy 641 IVIFLYISLALGSYSWSRVWDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVILQ 700
Db 403 LILMAYACVTM---LRWD---CAQSQSVGLAGVLVALAVASGLGCLALLGITFNAATTQ 457
Qy 701 VVPLVLVSGADNIFVLEVQRLPRRPGEPREHIGRALGRVAPSMILCSLSEALCFPL 760
Db 458 VLFFLALGIGVDDVFL--LAHFTALPOTPLQERNGECLQRTGTSVLTINNMAAFILM 515
Qy 761 GALTTPAVRTPALTASGLAVILDFLLQMSAFVALLSLQSKQESRLDVCC----- 812
Db 516 ALVPIPALRAFSLQAAIVVGCTFVAVMLVFPAILSLDLRRRCORLDLVCCFSFSCSAQ 575
Qy 813 ---VKPQEL-----PPQGGGGL--- 827
Db 576 VIQLPQELGDTVPVGIHAHLTATVQAFTHCBASSQHVVTILPPQAHLPVPPSDPLGSEL 635
Qy 828 -----LIG-----PFQKAYAFPLHMTIRGVVLLFLFA 855
Db 636 FPGGSTRLLQGBETROKAKKSLPCARWNLAFARVQFAPLQLQSHAKAIVLVLFGA 695
Qy 856 LFGVSLYSCHISVGLDQDELALPKOSYLLDYFLFNRYFEVGAAPVYFVTLGNFNS-SEA 914
Db 696 LGLSLSYGATLVQDGLATDVVPRGPKHAFLSAQLRYFSL-YEVALVTVGGFDYAHSQR 754
Qy 915 GMAICSSAGCWNFSFTQKIQATPEPEQSYLAIPAS-----SWDDFDIDLWLTSSCC--- 967

Db	870	-GLAASQANTYPPPPPEWLHKYDTTGENLRIPRAAQPLEPAQFPQFLLHGLQKTADFVEAIE	928
Qy	1068	AARELAANI -TADLRKVPGTDPFAVEFPYITINVFYEQYITILPEGLFMLSIC--LVPTF	1124
Db	929	GARAACTEAGQAGVHAYPSGSFF-----LFWEQYIGL--RRCFLLAVACILLVCTF	976
Qy	1125	AVSCLLIGLDLRSGLLNLLSTVMILVDTVGFMAWDISYNNAVSLINIVSAGMSVEFVSH	1184
Db	977	LVCALLLLSPWTAGLI-VLVLANMTVELFGIMFGIQLKLSAIPVWILVASIGIGVEFTVH	1035
Qy	1185	ITRSEFATSKETWLERAKEATISMGSAVFAGV---AMTNLPGLIIVGLAKAQLTQIEFFR	1241
Db	1036	VALGFLTSHGSRNLRAA-----SALEQTFAPVTDGANVSTLLGLLLAGSNPDFIIRFFV	1090
Qy	1242	LNMLITLLGLLHGLVFLPVILSYGVP	1267
Db	1091	VLVTLLTLLGLLHGLLLPVLISLIGP	1116

RESULT 59	
AAAY92703	
ID	AAAY92703 standard; protein; 1203 AA.
XX	
AC	AAAY92703;
XX	
DT	10-AUG-2000 (first entry)
XX	
DE	Human patched 2 (PTCH2) protein splice variant.
XX	
KW	Patched 2; PTCH2; lp32-35; tumour suppressor; familial melanoma CMM1;
KW	familial adenomatous polyposis; Momi; Michelin Tire Baby Syndrome;
KW	sonic hedgehog; SHH; development; carcinogenesis; cytostatic;
KW	basal cell carcinoma; gene therapy; chromosome lp32-35; splice variant.
XX	
OS	Homo sapiens.
XX	
FN	WO200020037-A1.
XX	
PD	13-APR-2000.
XX	
PF	06-OCT-1999; 99WO-SE001784.
XX	
PR	06-OCT-1998; 98SE-00003393.
XX	
PA	{KARO-} KAROLINSKA INNOVATIONS AB.
PA	{DHAA } PHARMACIA & UPJOHN CO.

XX Zaphiropoulos PG, Unden AB, Toftgard R, Rahnama F;
PI Hollingsworth RE;
XX WPI: 2000-303645/26.
XX N-PSDB; AAA09084.
XX Isolated human protein capable of participating in human patched
PT gene/sonic hedgehog pathway during embryonic development. is used in
PT medicament for treatment of condition involving tumors such as basal cell
PT carcinoma.

xx Disclosure; Page; 55pp; English.

xx This protein is a splice variant encoded by the novel human patched 2

xx gene (PTCH2), which has been localised by radiation hybrid mapping to

xx chromosome 1p32-35 with DIS211 and W11404 as closest flanking markers and

xx with an estimated localisation 5.5 cR from DIS443. This region is often

xx lost by LOH in various different tumour types, such as neuroblastoma,

xx melanoma and breast and colon cancer. PTCH2 is a candidate for a tumour

xx suppressor gene in this region. It is also a candidate gene for

xx involvement in familial melanoma CMM1, modifier locus for familial

xx adenomatous polyposis hMOM1 and Michelin Tire Baby Syndrome. PTCH2 is

xx capable of participating in the human patched gene/Sonic hedgehog

xx (PTCH/SHH) pathway during embryonic development and/or carcinogenesis.

xx The isolated human protein is useful as a medicament for the treatment of

xx a condition involving tumours such as BCC (basal cell carcinoma). The

CC nucleic acid is useful in gene therapy, and for use as a probe, primer or
CC a diagnostic agent. Note: This sequence was constructed using information
CC from Figure 2B and the protein sequence given on pages 41-44 of the
CC specification

	Query Match	8.4%; Score 583; DB 3; Length 1203;
	Best Local Similarity	22.9%; Pred. No. 4,7e-44;
	Matches	284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
Qy	306	RVAPARDKSMVDPKGTGTSLSDKL--SPSTHTLLGQFPQG-----WGTWVASWNPULITVLVS 359
Db	3	RSFPLRELPPSTTPPARTAAPQILAGSLKAPLWLRAYFQGLLFSLGGCIGRHCKGVFLG 62
Qy	360	VIPVVALAGLVFETLTDPVELMSAPNSQARSEKAFHQHFG--PPERTNOVILTAPNR 417
Db	63	LLAFGALALGLRMAIETNLEQLWVEVGSRSVQELHYHTEKLGLEAAVTSQMLQYARQE 122
Qy	418	SSVRYDSLGLGPNFSGILDLLLELELEQLERHLQVSPQEARQNSLQDICY--AP 474
Db	123	GEN-----LITPEALG-----LHLQAALTASKVQVSLYKSKMDLNKICVKSQVP 166
Qy	475	LNPDN-----TSLYDCCINSLLQVFNQNRITLLLTANQTLMGQTSQVDWKD----- 520
Db	167	LIENGMIERMIKLPFCVILTPDCCFWECAK---LQGSAYVLPGRPDIQWNLNDEQULLE 223
Qy	521	-----HFLYCANAPLTFKDGKTALALSCHMADYGA 548
Db	224	ELGFFASLEGFRELLEKQVQNYGVRPCPLHPDDHLCPSPANHHSRQAPNVAAHLSGCG 283
Qy	549	PVPFP-----LAIGGYKGDYSE---AEALIMTF-----SLNTPAGDPRLQAOKL 591
Db	284	HGFSHKFWHQEELLGGMARDPQGEELLRAEALQSTFLLMSPROLYEHRGDIYQTHDIGW 343
Qy	592	WEZAFLEMCAPQRRMAGHFOVTFTRERSLE--DEINRTTAEDLPATFS-----Y 640
Db	344	SEEQASTVLAQWRQRRFVQLAQEALPENASQOIHAFSSTITLDDI--LHAISEVSAARVGGY 402
Qy	641	IVIPLXISLALGSYSMSRVMVDISKATLGLGGVAVVLGAVMAAMGFFSYLGRSRSLVILQ 700
Db	403	LIMLACVYM---LRWD---CAQSQSGSVGLAGVLLVALAVASGLGLCALLGTFNAATTQ 457
Qy	701	VVPLVATSVGADNITPILVLYQRLPRRPGEPREHVHIGRALGRVAPSMLCISEALCPFL 760
Db	458	VLPLALGIGVDDVFL--LAHAFTEALPGTFLQRMGECLOQTGTSVVLSTINNNAAFLM 515
Qy	761	GALTMPARVTRFALTSGLAVILDFLLQMSAFVALLSDIKRQESASRLDCCC----- 812
Db	516	AALVPILPAIRAFSLQAAIVVGCTFVAVMLVFPAILSLDLRRHRCQRLDVLCCFSSPCSAQ 575
Qy	813	---VKPQEL-----PPQCGEGL-- 827
Db	576	VTIQLPQLGDTVPVGTIAHLTATVQAFTHCEASSQHVVTILPQPAHLVPPSPDPLGSEL 635
Qy	828	---ILG-----PFQKAYAPFLHLHWITRGVLLFLFLA 855
Db	636	PSFGSGTRDLGQBEETRQKAAKSKLPCARWNLAHFARYQFAPLQLQSHAKAIVLVFGA 695
Qy	856	LFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGPVYVVTVLGNFVS-SEA 914
Db	696	LLGLSLYGNLTVQDGLALTDDVPRGTKEHAFSLAQRLYPSL-YEVALVTQGGFDYAHSQR 754
Qy	915	GNNAICSSAGCNPFSFTQKIQYATEFPEQSYLAIPAS---SWDDDFIDMLTSPSSCC--- 967
Db	755	AL-----FDLHQRFSSL-----KAVLPPPTQAPRTMLHYIRYNLQGIQAAPDQ 798
Qy	968	-----RLYISGPNKDKFCPSTVNSLNCILKNCMSITMGSRVPSVEQFHKYLPMFLND 1018
Db	799	DWASGRITHTSTRNG-SED-----GALAYKLLIQTGAQELLD--FSQLTTRKLVD 846
Qy	1019	RPNIKCPKGLAAYSTSVNLTSDDGVLAASRFMAVHKP----- 1055

Db 847 REGLIPPE--LFYMGTLVWVSDPLGLAASQANFYPPPPPEWLHDKYDTTGTFNFRIPPAQP 904
 QY 1056 -----LKNQDYTEALRAARELAANI-TADLRKVPCTDPAFEVFPVTITNFFY 1102
 Db 905 LBPAPFPFLRLGLQKTADFVEALEGARACACAGQAGVHAYPSGSP-----LFW 954
 QY 1103 EOYLFTLPLGLFMLSIC--LVPTFAVSCLLGLDLRLSGLNLISIVMLVDTVGFNMLWD 1160
 Db 955 EOYLGIL--RRCFLAVCILLVCTFLVCALLLNPPWTAGLI-VLVLAMTVVELFGINGFLG 1011
 QY 1161 ISVNAVLNLNLSAVGMSVEFYVSHITRSFAISTKPTLWLRKAEATISMGSAVPAGVAMTN 1220
 Db 1012 IKLSAIPVILVASVGIGVEFTVHALGP-LITQSGNRRAHALEHTPVDG-AIST 1069
 QY 1221 LFCILVLGLAKAQILQIFFRNLNLITLGLLHGLVFLFVILSVYGVDPNPALEOKRA 1280
 Db 1070 LLGLMLAGSHDFIVRYFFAALTVLTLGLLHGLVLLPVLLSILGPP--PEVIQWKES 1127
 QY 1281 EBAVAV-----WVASCNHPHSVSTADNIYVN 1308
 Db 1128 PEILPAPPOGGGLRWGASSLPQSFARVTTSMTVAIH 1165

RESULT 60

AAV28444
 ID AAY28444 standard; protein; 1203 AA.

AC AAY28444;

DT 03-DEC-1999 (first entry)

XX Human ptc-2 protein.

Patched-2; ptc-2; human; hedgehog receptor; nootropic; neuroprotective;
 antiinflammatory; antiparkinsonian; cardiant; antiarthritic; screening;
 modulator; antagonist; agonist; cellular proliferation; neuronal tissue;
 testicular tissue; osteogenic tissue; chondrogenic tissue; disease;
 graft; transplant; treatment; nervous system injury; chemical injury;
 vasa injury; infection; inflammatory; tumor-induced injury; aging;
 Alzheimer's disease; chronic neurodegenerative disease; innervation;
 Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;
 spinocerebellar degeneration; multiple sclerosis; autonomic disorders;
 peripheral nervous system; smooth muscle; endocrine tissue; tachycardia;
 atrial cardiac arrhythmia; cell differentiation; chronic pain syndrome;
 lesion-induced death; neuron regeneration; damage repair; skeletal;
 cartilage; osteogenesis; arthritis; bone fracture; hereditary disease;
 prosthetic cartilage device; spermatogenesis; fertility enhancer.

XX Homo sapiens.

OS WO929854-A1.

PN 17-JUN-1999.

PD 08-DEC-1998; 98WO-05026009.

PF 08-DEC-1997; 97US-0067940P.

PR (ONTO-) ONTOGENY INC.

XX Bumcrot DA;

XX WPI; 1999-561298/47.

XX N-PSDB; AAX89478.

New human patched-2 (ptc-2) genes and proteins, useful in the treatment,
 prevention and/or reduction of the severity of neurological conditions.
 Claim 3; Page 73-77; 80pp; English.

This invention describes a novel recombinantly produced human patched-2
 (ptc-2) polypeptide which has nootropic, neuroprotective, cardiant,
 antiinflammatory, antiparkinsonian and antiarthritic activity. The ptc-2

CC protein is a hedgehog receptor and is therefore capable of modulating
 CC hedgehog signalling, and so affect a number of hedgehog-mediated
 CC biological activities. The human patched-2 (ptc-2) protein can be used to
 CC screen for modulators, antagonists and agonists, which are likely to play
 CC an important role in the modulation of cellular proliferation and
 CC maintenance of, e.g. neuronal, testicular, osteogenic or chondrogenic
 CC tissues during disease states. Modulators of ptc-2 protein can be used
 CC for in vivo reformation of tissue; to improve grafting and morphology of
 CC transplanted tissue; for the treatment, prevention and/or reduction of
 CC the severity of neurological conditions deriving from: injury to the
 CC nervous system including traumatic injury, chemical injury, vascular injury
 CC and deficits (such as ischemia resulting from stroke), together with
 CC infectious/inflammatory and tumor-induced injury; ageing of the nervous
 CC system including Alzheimer's disease; chronic neurodegenerative diseases
 CC of the nervous system including Parkinson's disease, Huntington's chorea,
 CC amyotrophic lateral sclerosis, as well as spinocerebellar degenerations;
 CC and chronic immunological diseases of the nervous system including
 CC multiple sclerosis. ptc-2 therapeutics can also be used in the treatment
 CC of autonomic disorders of the peripheral nervous system, including
 CC disorders affecting the innervation of smooth muscle and endocrine
 CC tissue, e.g. to treat tachycardia or atrial cardiac arrhythmias.
 CC Antagonists of ptc-2 protein can be used to prevent differentiation of
 CC cells in culture, as well as for treatment of chronic pain syndromes.
 CC Agonists may be used to rescue neurons from lesion-induced death as well
 CC as neuron regeneration, in diseases such as CNS trauma infarction, (viral)
 CC infection, metabolic disease, nutritional deficiency, toxic agents, and
 CC so on. ptc-2 therapeutics may also be used for the repair of central and
 CC peripheral nerve damage, for repair and regeneration of non-neuronal
 CC tissue, e.g. skeletal and cartilage tissue, e.g. in the treatment of
 CC osteogenesis, arthritis, bone fractures, hereditary disease, as well as
 CC for generation of prosthetic cartilage devices, and to induce
 CC spermatogenesis and as fertility enhancers. This sequence represents the
 CC human ptc-2 protein described in the invention

XX Sequence 1203 AA;

SQ Query Match 8.4%; Score 580; DB 2; Length 1203;

Best Local Similarity 22.9%; Pred. No. 8.9e-44;

Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

QY 306 RVAPARDKSKWDPKKGTSLSDKL--SPSTHTLLQOPFG-----WGTWVASHPLTILVLS 359
 Db 3 RSPPLRELPSPSTPPARTAAPOILAGSLKAPLWLRAYFQGLFSLGCGIQRHCKGVFLG 62
 QY 360 VIPVVALAAGLVFTLTLLTDPVELSAPNSQASRKAFTDHFQ--PFFRTNVOILLTAPNR 417
 Db 63 LIAFGALALGLRMAIETNLEQLWVEVGSVQELHYTKELGEBEAAYTSQMLIQTRQE 122
 QY 418 SSRYDSSLGPKNPSGILDLLELLELLELRLHQLVWSPQAQNTSLQDICY---AP 474
 Db 123 GEN-----ILTPALG-----LHLQAALTASKVQVSLYKGSWDLNKICYKSGVP 166
 QY 475 LNPDN-----TSLYDCCINSLLQVFNNTTLLLTANGTLMGQTSQVDWKD----- 520
 Db 167 LIENGMIERMTEKLPFCVILTPDCFWGAK---LQGGSAVLPGRPDQMTNLDPEQLLE 223
 QY 521 -----HFLYCANAPLTFKDGATLALSCMADYCA 548
 Db 224 ELGPFASLEGFRELDDKAQVQYVGRPCLPHPDDLHCPPSAPNHHSRQAPNVHLSGGC 283
 QY 549 PVVPPF-----LAIGGYKGDYSE---AEALIMTF-----SLNNTYPAGDPRLAAKL 591
 Db 284 HGFSHKPMHMQBELLLGGWARDPQCELLRABALQSTFLLMSPRQVYHFRGQYQTHDIGW 343
 QY 592 WEEAFLEENRAPORRMAGMFQVTFATRSLE--DEINRTTAEDLPFATS-----Y-640
 Db 344 SEEQASTVLOAQORRFVOLAQALPENASQQTHAFSSTTLDI-LHAFSEVSAARVWGY 402
 QY 641 IVIFLYISIALGSSYSSWRVWVDSKATLGLGVAVVLGAVMAAMGFFSVLRSSILVILQ 700
 Db 403 LLMLAYACTM---LRWD--CAQSGSVGLAGVLLVALAVASGLGLCALGTLFTNNAATQ 457
 QY 701 VVPFLVLSVGADNIFVLEVQLPRRPPCEPREVHIGRALGRVAPSMILLCSLSEATCFPL 760

```

Db      458  VLPFLALGIGVDDVFL--LAHAFTALPGTLOERMGECLQRTGTSVLTSTINMAAFLM 515
QY      761  GALTMPAVRTFALTSLGLAVILDLQMSAFVALLSLDSKROEASRLDVCCE----- 812
Db      516  AALVFPALRAFSQAALVVGCTFVAVMLVFPALFSLDLRRHCRDLVLCFSPSCSAQ 575
QY      813  ---VKPQBL-----PPPGGEGE-- 827
Db      576  VIQILPQELGDTVPVGTIAHLTATVQAFTHCEASSQHVVTILPQAHLVPPSPDLGSEL 635
QY      828  -----LLG-----FFQKAYAPFLHWTIRGVLLFLA 855
Db      636  FSPGSTRDLGQEBEETRQKACKSLPCARWNLAHFARYQAPFALLOSHAKAIVLVFGA 695
QY      856  LFGVSLYSNCHTSVGLDQELAPKDSYLLDPLFLNRYFEVGAPVYFVTILGYNFS-SEA 914
Db      696  LGLSLYGATLVQDGLALTDDVPRGTKEHAFLSAQRLYFSL-YEVALVTQGGFDVAHSQR 754
QY      915  GMAAICSSAGCMNFSFTQIOVATEFPBQSYLAIPAS-----SWDDFDIMLTPSSCC--- 967
Db      755  AL-----FDLHQRFSSL-----KAVLPPPATQAPRTWLHYRNWLQIGIQAAFDQ 798
QY      968  -----RLYISGNKDKPCPSTVNSLNCILKNCMSITMGSVRSVQFHKYLPWFELND 1018
Db      799  DWASGRITRHSYRNG-SED-----GALAYKLLIQTGAQBPLO-FSOLTRKLVD 846
QY      1019  RPNIKCPKGLAAYSTVNLTSIDQVLAASFMAHYKP-----1055
Db      847  REGLIPPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQP 904
QY      1056  -----LKNSQDYTEALRAARELAANI-TADLRKVGTDPAPAEVFPYTTINVFY 1102
Db      905  LEPAQPFLLRLGLQKTADFVEATEGARACAEAGQAGVHAYPSGSPF-----LFW 954
QY      1103  EQVLTILPGLMLSLC--LVPTFAVSCLLGLDLSRLNLSIVMLVDTVGFPMALWD 1160
Db      955  EQVLGL--RRCFLAVCILVCTFLVALLNPNWAGLI-VLVLAAMTVELFGIMGFLG 1011
QY      1161  ISYNAVSLINLVSAGVMSVEFVSHITRFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
Db      1012  IKLSAIPVVLVASGVIGVEFTVHVALGF-LTTQGSRLRAAAHLEHTAPVTDG-AIST 1069
QY      1221  LPGLVLGLAKAQIQLFPRMLNLLTLGLLHGLVFLPVILSYVGVDPVNPALALEOKBA 1280
Db      1070  LLGLLMLAGSHFDFIVRYFPAALTVLTLGLLHGLVLLPVLSILGPP--PEVIOMYKES 1127
QY      1281  BEAFAV-----MVASCPNHPRSVSTADNIYVN 1308
Db      1128  PEILSPAPOGGGLRWGASSLFSQSFARVTTSTMTVAIH 1165

RESULT 61
RAY92225
ID  AAY92225 standard; protein; 1146 AA.
XX
AC  AAY92225;
XX
DT  10-AUG-2000 (first entry)
XX
DE  Human patched 2 (PTCH2) protein.
XX
KW  Patched 2; PTCH2; tumour suppressor; familial melanoma CMM1;
KW  familial adenomatous polyposis; hMomi; Michelin tire Baby Syndrome;
KW  sonic hedgehog; SHH; development; carcinogenesis; cycostatic;
KW  basal cell carcinoma; chromosome 1p32-35; gene therapy.
XX
OS  Homo sapiens.
XX
FN  WO200020037-A1.
XX
PD  13-APR-2000.
XX

```

```

PF 06-OCT-1999; 99WO-SE001784.
XX
PR 06-OCT-1998; 98SE-00003393.
XX
PA (KARO-) KAROLINSKA INNOVATIONS AB.
PA (PHAA-) PHARMACIA & UPJOHN CO.
XX
PI Zaphiropoulos PG, Unden AB, Toftgard R, Rahnama F;
PI Hollingsworth RE;
XX
DR WPI; 2000-303645/26.
DR N-PSDB; AAA09081, AAA09084.
XX
PT Isolated human protein capable of participating in human patched
PT gene/sonic hedgehog pathway during embryonic development is used in
PT medicament for treatment of condition involving tumors such as basal cell
PT carcinoma.
XX
PS Claim 2; Page 41-44; 55pp; English.
XX
CC This protein is encoded by the novel human patched 2 gene (PTCH2), which
CC has been localised by radiation hybrid mapping to chromosome 1p32-35 with
CC DIS211 and W1404 as closest flanking markers and with an estimated
CC localisation 5.5 cR from DIS443. This region is often lost by LOH in
CC various different tumour types, such as neuroblastoma, melanoma and
CC breast and colon cancer. PTCH2 is a candidate gene for involvement in
CC gene in this region. It is also a candidate gene for familial adenomatous polyposis
CC familial melanoma CMM1, modifier locus for familial adenomatous polyposis
CC hMomi and Michelin Tire Baby Syndrome. Splice variants of the PTCH2 gene
CC exist (see AAA09082-84). PTCH2 is capable of participating in the human
CC patched gene/sonic hedgehog (PTCH/SHH) pathway during embryonic
CC development and/or carcinogenesis. The isolated human protein is useful
CC as a medicament for the treatment of a condition involving tumours such
CC as BCC (basal cell carcinoma). The nucleic acid is useful in gene
CC therapy, and for use as a probe, primer or a diagnostic agent
XX
SQ Sequence 1146 AA;
Query Match 8.4%; Score 579; DB 3; Length 1146;
Best Local Similarity 23.3%; Pred. No. 1e-43;
Matches 276; Conservative 161; Mismatches 452; Indels 298; Gaps 40;
QY 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLQCFQPG-----WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSTTPARTAAPQILAGSLKAPLWRAVFGLLFSLGCGIQRHCGKVLFLG 62
QY 360 VIPVVALAAGLVFTTLTDPVELMSAPNSQARSEKAFHDQHF--PFRTNOVILTAENR 417
Db 63 LIAFGALALGLRMAIETNLSQLWVEVGSRSVQELHYTEKLGEEAAYTSQMLIQTAQGE 122
QY 418 SSVRYDSLILGPKNFSGLDLDLLELLELQERLHLQVMSPEAQORNISLODICY---AP 474
Db 123 GEN-----LHQAALTASKVQVSLYKGSWDLNKKYKSGVP 166
QY 475 LNPDN-----TSLYDCCINSLQVQNNRTLLLTANQTLMGQTSQVMDKD----- 520
Db 167 LIENGMIERMIEKLPFCVILTDLDFWEGAK---LQGSAYLPGRPDIQNTWLDPEQLLE 223
QY 521 -----HFLYCANAPLTFKDGKTALALSCMADYGA 548
Db 224 ELGPASLEGFRELIDKAQVGYVGRPCPLHPDDLHCPPSAPNHHSRQAPNVAHELSGGC 283
QY 549 PVPPPP-----LAIGYKGYKDYSE---AEALIMTF-----SLNNYPAGDPRLAQAKL 591
Db 284 HGFSHKFMHWQEBELLIGGMARDPQGBELLRAELAQSTFLMSPRQLYEHFRGDIYQTHDIGW 343
QY 592 WEEAFLEEMRAFORRMAGMFQVTTFAERSLE-DEINRTTAEDLPITPATS-----Y 640
Db 344 SEEQASTVIOAWORRFVQLAQAEALPENASQQIHAFSSITLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISIALGASYSWSRVMVDSKATLGLGGVAVVYLGAVMAMGFSSVILGRSSVLILQ 700
Db 403 LLMLAVACTVM---LRWD--CAQSQSGSVGLAGVLLVALAVASGLGICALLGITFTNAATTQ 457

```


Db 528 P-----FSLDVRNEDETOQIASREFLIQAVNITDTHHEKEMVR----- 565
QY 1079 DLRKVPGETD--AFEVFPYTTITNVEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLR 1136
Db 566 DLRQICKDPSLNASTFHPYP---VFQDFELVRPVSQAMVIGAILMMIISFVIP-NIL 621
QY 1137 SGLNLLSIVMLVDTGFMALWDISYNVSLINLVSAGMSVFEVSHITSPALSTKPT 1196
Db 622 CSLWAFVSISTELGVAGYMALWDVNLDSISMINIMCIGFSVDFTAHICYTMSKKRS 681
QY 1197 WLERAKEATINGSVAFFAGVAMNLPGLTIVLGLAKAQLIQIFFPRLNLLITLLGLLGLV 1256
Db 682 PKARREALHSLUGLPIIOGSSTIL-GIVALLLAQSYIP-LVFFKQVFLVIFFGAMHGLF 739
QY 1257 FLPVILSYVGP 1267
Db 740 LLPVLLSLFGP 750
RESULT 63
ID AAW72971
AAW72971 standard; protein; 1311 AA.
XX AAW72971;
XX
DT 17-OCT-2003 (revised)
DT 26-JAN-1999 (first entry)
XX
XX Precis coenia patched amino acid sequence.
XX Patched gene; ptc; diagnosis; treatment; developmental disorder; cancer;
KW healing; injured tissue; spina bifida; Wnt-1 oncogene; sperm production;
KW gene therapy.
XX Junonia coenia.
XX
FH Key Location/Qualifiers
FT Misc-difference 348
FT /note= "unspecified"
FT Misc-difference 908
FT /note= "unspecified"
XX
FN US5837538-A.
XX
PD 17-NOV-1998.
XX
PF 06-OCT-1995; 95US-00540406.
XX
PR 07-OCT-1994; 94US-00319745.
XX
PA (STRD) UNIV LELAND STANFORD JUNIOR.
XX
PI Scott MP, Goodrich LV, Johnson RL;
XX
DR WPI; 1999-023461/02.
XX
XX Nucleic acid encoding vertebrate patched protein and related
PT transformants - used to express poly:peptide(s), useful for diagnosis and
PT treatment of developmental disorders or cancer, and in healing of injured
PT tissue.
XX
XX Example; Col 25-32; 38pp; English.
XX
XX The present invention describes vertebrate and invertebrate patched (ptc)
XX genes. Cells containing and expressing the ptc gene are used for the
CC recombinant production of the protein. These in turn are useful: (i) for
CC generating antibodies (Ab); and (ii) to screen for specific-binding
CC ligands (potential therapeutic agonists and antagonists). The ptc gene,
CC or its fragments, are used to isolate related sequences from other
CC mammals; to identify mutations (particularly those associated with
CC genetic diseases such as spina bifida and other developmental disorders);
CC to monitor expression levels in testis (to determine relationship with
CC sperm production) and to isolate 5'-non-coding sequences (used to study

embryonic development and to provide regulated expression of proteins).
The complete gene can be used in gene therapy, including expression of
antisense molecules, and to generate transgenic animals for studies of
embryonic development. Ab are used diagnostically to determine the ptc
protein on cell surfaces and as competitive inhibitors of signal
transduction through the ptc ligand. Cells that have been engineered to
express the ptc protein can be used to promote regrowth and healing of
damaged tissue (e.g. growth of new teeth) and regulation of the ptc
protein expression may be useful in cancer treatment (it may control the
Wnt-1 oncogene). The present sequence represents Precis coenia
(butterfly) patched amino acid sequence, from the present invention.
(Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1311 AA;
Query Match 7.6%; Score 526.5; DB 2; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
QY 336 ILGQFFQGGTWTWASWPLTILVSVIPVWVLAAGLVFTLTTPVELMSAPNSQARSEKA 395
Db 72 ILGCFLOGDAGKVL F--VAILVLTFCV-----GLKSAQIHTRVDQLWVQSGRLAEALK 124
QY 396 FHDQHFQFP-FRTNQVILTAENRSYRYDSLLGLPKNFSGILDLDLLELLELQRLRLH 454
Db 125 YTAQALGEADSSHTQLVI---QTAQKPDVSLHP---GAL-----LEHL 162
QY 455 QVWSPQARNI-----SLQDICVAPLNP-----NTSLYD-----CCINSLLQYFQNN 497
Db 163 KVVHAATRVTVHMYDIEWRLKDLCSYSPIDPFEGYHHIESIDNVIPCAIITPLDCFWEG 222
QY 498 RTLL-----LITANQTLMGQTSQ--- 515
Db 223 SKLGLPDYPIVPHLKHKLQWHLNPLEVVEVKLKFQFPPLSTIEAYMKRAGITSAYMK 282
QY 516 ---VDWKDHFLYC-ANAPLTFKDGFTALALSCWADYGAPVFPF-----LAIGG----- 558
Db 283 KPCLDPTDP--HCPATAP-NKSGHI PDVAAELSHGCGFAAAVMHMQEQLITVGATRNS 339
QY 559 ----YKGDYSEAEALIMTFSLNYPAGDPRLAQAKLMEAEFLMEARFORFMAGMFQVT 614
Db 340 TSALEKARKLQTVVQLMGEREMYEYADHYKVHQIGWNOEKAAAVLDQWQKFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPIFATSVIVIFLYISIALGSYSSWSRYM 661
Db 400 TTSGSVSAYSPYFPSTSTINDILGKFSVSLKNILGYMFMLIYVAVTL---IQW-RDP 455
QY 662 VDSKATLGLGVAVVVGAVMAAGPFYSYLGRSSILVILQVVPFLVLSVGADNIFIFVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASTQIVPFLALGLGVQDMFLTHTY 515
QY 722 QRLPRRPGEPREVHIGRALGRVAPSMILCSLSEACFFLGLGALTPMPAVRTEALTSLGLV 781
Db 516 --VEQAGDVPREERTGLVLKKSGLSVLLASLNVAFALAAALPIPAFRVFLQAAIILL 573
QY 782 LDPLQLMSAFVALLSLDSKQBSRLDVCCKVKPQELPP----- 820
Db 574 FNLGSILLVFPAMISLDLRRRAARADLLCCLMPESPFPKKKIPERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGEGL-----LLGFFQKAYAPFLHLHITRGVVLVLLFLALFVSLYSMC 865
Db 634 TRQPLDPDVSENVTCTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTYSWGAT 693
QY 866 HLSVGLDQELALPKDSYLLDYFLNLR---YFEVGAPVYFVTTLGNFSSSEAGMNAICSS 922
Db 694 KYVQGLDLTDIPENT---DEHEFUSRQEKYFGF-----YN-----MYAVTQ- 732
QY 923 AGCNFNSF--TKIQYATEFPQSYLAIP-----ASSWVDDFDWL----- 961
Db 733 ---GNFEYTNKLLY--EYHDQ-FVRIENIKNDNGGLTKFWLSLFRDWLLDLQVAFDK 786
QY 962 -TFSSCCRLYISGPNKDKFCPSPTVSNLCKMCKMSTMGSVRPSVEQ----- 1007

Db 787 EVASGCI-----TQEWCKNASDE--GILAYKLMVQTGHVDNPIDKSLITAGHRLVXD 838
 Qy 1008 -----FHKYL--PWFINDRPNKCPKGGI-----AAYSTVNLTSDGQVLASRFM 1050
 Db 839 GIINPRAFTNYLSAWATNDALAYGASQGNLKPOQRWHSPEVDHLEIKKSSPLIYTQLP 898
 Qy 1051 AYHKPLKNSODYTEALRAARELANITADLRKVPGTDPAPFPPYITITNVEQYLTILP 1110
 Db 899 FYLSGLSDTXSIXTKILRSVRDLCKYEA--KGLPN-----PPSGIPFLFWQYLYLRT 949
 Qy 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLINLLSIVMILVDVTGFMALWDISYNVSLIN 1170
 Db 950 SLLALACALAAVF--TAVWVLLNAAVAALVTLALATLVQLLGVNALLGVLSAPAVL 1008
 Qy 1171 LVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAVAMTNLPGLVGLA 1230
 Db 1009 LVLAIGRGVHTVHLCLGFTSICG---KRRASLALES--VLAPV---VHGALAAALA 1059
 Qy 1231 KQAL-----IQIFFRMLNLTILGLLHGLVFLPVILSYGP--DVNPALALEOKRAB 1281
 Db 1060 ASMLAASECGFVARLFLRLDLDIVFLGLIDGLLFFPILVLSILGFAAEVRPI-----1110
 Qy 1282 EAAVAVMVASCPNHRKRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 64

AAB67156
 ID AAB67156 standard; protein; 1311 AA.

XX AAB67156;

12-APR-2001 (first entry)

XX Butterfly patched protein.

XX Butterfly; patched; PTC; segment polarity; limb patterning; development;

XX hedgehog; antibody; human; mouse.

XX Junonia coenia.

XX US6172200-B1.

XX 09-JAN-2001.

XX 20-OCT-1997; 97US-00954668.

XX 07-OCT-1994; 94US-00319745.

XX 06-OCT-1995; 95US-00540406.

XX (STRD) UNIV LELAND S STANFORD.

XX Scott MP, Goodrich LV, Johnson RL;

XX WPI; 2001-136884/14.

XX N-PSDB; AAF32177.

XX Novel monoclonal antibody useful in diagnostic assays for detection of

XX presence of protein on surface of cells specifically binds to naturally

XX occurring patched protein, other than Drosophila patched protein.

XX Disclosure; Col 23-30; 39pp; English.

XX The present invention provides a monoclonal antibody which specifically

XX binds to a patched protein (PTC) other than that from Drosophila. Also,

XX given are the protein and coding sequences of patched from the beetle,

XX mouse, butterfly, fruit fly, mosquito and human. Patched is a segment

XX polarity gene involved in limb patterning. The sequences can be used to

XX study development and to isolate the patched ligand, hedgehog. In

XX addition, antibodies can be used to detect the PTC protein on cell

XX surfaces or to inhibit the transduction of signal by the PTC ligand by

XX competing for its binding site

XX SQ Sequence 1311 AA;
 Query Match 7.6%; Score 526.5; DB 4; Length 1311;
 Best Local Similarity 22.2%; Pred. No. 1.1e-38; Indels 307; Gaps 46;
 Matches 258; Conservative 175; Mismatches 420;
 Qy 336 LLGQFPQGMGTWASWPLTILVLSVIPVVALAAGLVTELTTPDELWSAPNSQAKSEKA 395
 Db 72 ILGCFIQGAGKVLFF--VAILVLSFVCV-----GLKSAQIHTRVDQLWQEGGRLEAELK 124
 Qy 396 FHDQHGPPF--FRTNQVILTAPNRSYRYSLLIGPKNFGSILDLLELLELLELLELRLHL 454
 Db 125 YTAQALGEADSSHTQLVI---QTAKDQDVSLIHP---GAL-----LEHL 162
 Qy 455 QWSPAEQRNI-----SLQDICVAPLNPD-----NTSLYD-----CCINSILQYQNN 497
 Db 163 KVHAATRVTVMYDIEWRLKOLCYSPSPDFEGYHHISIIDNVPCAILITPDCFWEG 232
 Qy 498 RTLL-----LLTANOTLMQOTSO---515
 Db 223 SKLGGPDYPIYVPHLKHKLQWTHLNPLEVVEVKLKFQFPLSTIEAYMKRAGITSAYMK 282
 Qy 516 ---VDWKHFLYC--ANAPLTFKDGITALALSCWADYCAPVPPF-----LAIGG-----558
 Db 283 KPCLDPTDP--HCPATAP--NKESGHPDVAABELSHGCGYFAAAYMHWPEQLIVGATRNS 339
 Qy 559 ---YKGDYSEABALIMTFSLNNYPAGDPRLAQAQKLEAEPLSEMRAPQRRMAGMFPQVT 614
 Db 340 TSALRKARXLQTVVQLMGERMYEWADHYKHVQIGWNOEKAADAVDAMQKFAAEVRKI 399
 Qy 615 FT-----AERSLEDEINRTTAEDLPIPATSVIVIPLYISLALGSYSSRSRVN 661
 Db 400 TTSGSVSSAYSFPFSTSTLNDILGKFEVSLKXIIILGTFWFLIYVAVL---IQW-RDP 455
 Qy 662 VDSKATLGLGVAVVGLVAVMAAGPFSYLGIRSSVLQVPPVLVLSVGDNIPFIVLEY 721
 Db 456 IRSQAGVGIAGVLLLSITVAAGLGFALLGIPFNASSTQIVPELALGLGVQDMFLTHTY 515
 Qy 722 QRLPRPGEPREVIHGRALGRVAPSMILCSLSEACIFPLGALTTPMNAVTRFALTSLGLAVI 781
 Db 516 ---VQAGDVPREERTGLVLKKSGLSVLLASLQNVMAFLAAALLPIPAFRVFCQAAILLL 573
 Qy 782 LDPLQLMSAFVALLSLDSKROBASRLDVCCKVKPQELPP-----820
 Db 574 FNLGSLILVFPAMISLDLRRSARADLCLCLMPESPPLPKKIPERAKTRKDKTHRIDT 633
 Qy 821 -----PGQEGE-----LLGFFQKAYAPFLHWTIRGVVLLFLFALFGVLSYMC 865
 Db 634 TRQPLDPDVSENVTKCCLSVSLTKWAKQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
 Qy 866 HISVGLDQELALPKDSYLLDYFLNLR---YFVGVGAPVYFVTTLGNFSEAGMNAICSS 922
 Db 694 KVKQGLDLTDIVPENT---DEHEFLSRQEKYFGF-----YN-----MYAVTQ- 732
 Qy 923 AGCNPFP--TOKIQVATEFPEQSULAIR-----ASSWVDDFIDLW-----961
 Db 733 ---GNFEYTNQKLLY--EYHQDQ--FVRIENIINKDNGGLTKFWLSLFRDILLDLQVAFDK 786
 Qy 962 -TPSSCCRLYISGPNKDKFCPSVTNSLCLKNCMSITMGSVRPSVEQ-----1007
 Db 787 EVASGCI-----TQEWCKNASDE--GILAYKLMVQTGHVDNPIDKSLITAGHRLVXD 838
 Qy 1008 -----FHKYL--PWFINDRPNKCPKGGI-----AAYSTVNLTSDGQVLASRFM 1050
 Db 839 GIINPRAFTNYLSAWATNDALAYGASQGNLKPOQRWHSPEVDHLEIKKSSPLIYTQLP 898
 Qy 1051 AYHKPLKNSODYTEALRAARELANITADLRKVPGTDPAPFPPYITITNVEQYLTILP 1110
 Db 899 FYLSGLSDTXSIXTKILRSVRDLCKYEA--KGLPN-----PPSGIPFLFWQYLYLRT 949
 Qy 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLINLLSIVMILVDVTGFMALWDISYNVSLIN 1170

Db 950 SLLALACAAVE-IAVMVLLNNAWAVLTLATLVLQLLGVALLGVKLSAMPVL 1008
Qy 1171 LVSAVGSVEFVSHITSFALSTPTWLERAKETISNGSAVFAGVAMTNLPGILVLGLA 1230
Db 1009 LVLAIGRVHFTVHLCLGFTVSIGC---KRRRASLALES-VLAPV---VHGALAAALA 1059
Qy 1231 KAQL-----IQIFFPRLNLLITLGLLHGLVLPVILSYVGP--DVNPALALEQKRAE 1281
Db 1060 ASMLAASECGFVARLFLRLDLDIVPLGLIDGLLFFPIVLSILGPAAEVRPI----- 1110
Qy 1282 EAVAVMVASCPNHPSTVST 1301
Db 1111 -----EHERLST 1118

RESULT 65
AAG79573
ID AAG79573 standard; protein; 1311 AA.

AC AAG79573;
XX 29-AUG-2003 (revised)
DT 23-DEC-2002 (first entry)
XX Butterfly patched homolog.
KW Patched; ptc; basal cell nevus syndrome; BCNS; wound healing;
KW developmental disorder; spina bifida; craniofacial abnormality;
KW basal cell carcinoma; skin; brain tumour; transgenic; mouse;
KW proliferation; oncogenesis; signal transduction; cancer; aging.
XX Junonia coenia.

XX Key Location/Qualifiers
FH Misc-difference 348 /label= Unknown
FT FT Misc-difference 908 /label= Unknown
FT FT
XX US6429354-B1.
XX 06-AUG-2002.
XX 22-AUG-1997; 97US-00918658.
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX (REGC) UNIV CALIFORNIA.

XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX WPI; 2002-722086/78.
XX N-PSDB; ABA00376.
XX Transgenic mouse, useful for cancer drug screening, comprises disruption
XX of allele of patched locus which predisposes the mouse to develop a
XX proliferative disorder.
XX Example; Col 47-54; 58pp; English.
XX This sequence shows the butterfly patched (ptc) protein homolog.
XX Mutations in the patched gene are characteristic in basal cell nevus
XX syndrome (BCNS), an inherited disorder with an increased risk of
XX developmental disorders such as spina bifida and craniofacial
XX abnormalities, basal cell carcinoma of the skin and brain tumours in
XX mammals. The human and mouse ptc sequences show 86% homology. The
XX transgenic mouse of the invention has a genome comprising a disruption of
XX at least one allele of a patched locus, where the disruption predisposes
XX the mouse to develop a proliferative disorder. The transgenic mouse is
XX useful for studying patched function and regulation, for e.g. a series of

CC small deletions and/or substitutions made in the patched gene, to
CC determine the role of different exons in oncogenesis, signal
CC transduction, etc; and for drug screening which is useful for treating
CC cancer or developmental abnormalities attributed to a defect in patched
CC function such as wound healing and aging. Note: The specification
CC indicates that this sequence is encoded by the cDNA given in ABA00376.
CC However no ORF which decoded to the corresponding protein could be found
CC in either orientation. (Updated on 29-AUG-2003 to standardise OS field)
XX
SQ Sequence 1311 AA;

Query Match 7.6%; Score 526.5; DB 5; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
Qy 336 LLGQFFQGWGTWVASWPLTILVSVIPVVALAAGLVFTLTTPVELWSNPQARSEKA 395
Db 72 ILGCFLLQDAGKVLV--VAIVLSTFCV-----GLKSAQIHTRVQDLMWVGGRLEAEK 124
Qy 396 FHDQHFQFP-PRTNQVLTAPNRSYRYDSLLGPKNFSGILDLDLLELLELLELQERLHL 454
Db 125 YTAQALGEADSTHQLVI---QTAKDPDSSLHP-----GAL-----LEHL 162
Qy 455 QVMSPEAQRNI-----SLQDICYAPLNP-----NTSLYD-----CCINSLLQYQNN 497
Db 163 KVVAATRVVMYDIEWRLKDLCSPSIPDFEGYHHIESIIDNVIPCAITPLDCFEWEG 222
Qy 498 RTLL-----LLTANQTLMGOTSQ--- 515
Db 223 SKLLGPDYPIYVPHLKHKLQWHLNPLEVVEVKLKFQPLPSTIEATYMKRAGITSAVMK 282
Qy 516 ---VDMKDHFLYC--ANAPLTFKQGTALALSCMADYGAPVFPF-----LAIGG----- 558
Db 283 KPCLDPTDP--HCPATAP--NKSGHI PDVAELSHGCGYFAAAYMHMPEQLIVGATENS 339
Qy 559 ---YKGDYSEAEALIMTFSNNYPAGDPRLAQAKLWEAEFLBEMRAFQRMAGMFQVT 614
Db 340 TSALRKARXLQTVVQLMGEREMVEYWDHYKHVQHGMNQEKAAAVLDAMQKFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSWSRV 661
Db 400 TTSGSVSSAYSFPFSTSTLNDILGKFSEVSLKNILGYMFMILYIVATL---IQW-RDP 455
Qy 662 VDSKATLGLGVAVVGLGAVMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFVLEY 721
Db 456 IRSQAGVIGVLLSITVAAGLGFALLGIPFNASSQIVPFLALGLGVQDMFLTHY 515
Qy 722 QRLPRRGPBPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVTFAITSLGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGLSGLSVLASCNVMAFLAALLPIPAFRVFCLOAAILGL 573
Qy 782 LDFLLQMSAFVALLSLDSKQEAASLDVCCCVKQPELPP----- 820
Db 574 FNLGSILLVFPAMISLDLRRRAARADLLCCLMPESPPLPKKI PERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGE-----LLGFFQKAYAPFLLHMITRGVILLFLLFALGVSLYSWC 865
Db 634 TRQPLDPDVSENVTKTCCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
Qy 866 HISVGLDQBELALPKDSYLLDYFLNLR---YFEVGAPYFVVTTLGVNFSSEAGMNAICSS 922
Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYGF-----YN-----MYAVTQ- 732
Qy 923 AGCNNEFP--TOKIOVATFPQSYLAIP-----ASSWVDDFIDWL----- 961
Db 733 ---GNPEYPTNKKLY--EYHQ--FVRIINIKNDNGGLTKFWLSLFRDWLLDLQVAFDK 786
Qy 962 -TPSSCCRLYISGPNKDKFCPSTVNSLNCNKMCSITMGSVSRPSVQ----- 1007
Db 787 EVASGCI-----TQEYWCNKAASDE--GILAYKLMVQTGHVDNPDIDKSLITAGHRLVKD 838
Qy 1008 -----FKYL--PWFLNDRPNKCPKGGI-----AAVSTSVNLTSDDGOVLASRFM 1050

Db 839 GIINPKAFYNYLSAWATNDALAYGASQGNLKPQQRWIIHSPEDVHLEIKKSSPLIYTQIP 898
 QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEPFVPTITNVFYEQYLITLP 1110
 Db 899 FYLSGLSDTKSITKILRSVRDLCKYEA--KGLPN-----FPSGIPPLFWEQYLYLRT 949
 QY 1111 EGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMILVDTVGFWALWDISYNVSLIN 1170
 Db 950 SULLALACALAAVF-IAVMVLLNAAVAVLTALATLVLLQLLGVNALLGVLSAMPAYL 1008
 QY 1171 LVSAVMSVEFVSHITRSFAISTKPTWLRRAKEATISMGSAVPAGVAMTNLPGILVIGLA 1230
 Db 1009 LVLAIRGRGVHFTVHLCLGFVTSIGC-----KRRASLALES-VLAPV-----VHGALAAALA 1059
 QY 1231 KAQL-----IQIPFRRLNLLITLGLLHGLVFLFVLISYVGP--DVNPALALEQKRAE 1281
 Db 1060 ASMLAASECGFVARLFLRLLLDIVFLGLDGLLFFPFLVLSILGPAEVRPI-----1110
 QY 1282 EAAVAVMWASCPNHSRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 66

ABUG2149

ID ABUG2149 standard; protein; 1311 AA.

XX AC

XX ABUG2149;

XX 25-AUG-2003 (first entry)

XX DE

XX Buckeye patched gene PTC product.

XX KW

XX Buckeye; patched gene; PTC; hedgehog protein; gene therapy.

XX OS

XX Junonia coenia.

XX FH

XX Key Location/Qualifiers

XX FT Misc-difference 348

XX FT /label= Unknown

XX FT Misc-difference 908

XX FT /label= Unknown

XX XX

XX US2003032085-A1.

XX PN

XX 13-FEB-2003.

XX PD

XX PF 20-OCT-1997; 97US-00954701.

XX PR

XX 07-OCT-1994; 94US-00319745.

XX PR 06-OCT-1995; 95US-00540406.

XX XX

XX PA (SCOTT/) SCOTT M P.

XX PA (GOOD/) GOODRICH L V.

XX PA (JOHN/) JOHNSON R L.

XX XX

XX PI Scott MP, Goodrich LV, Johnson RL;

XX XX

XX WPI; 2003-492065/46.

XX DR

New DNA sequence other than present in a chromosome encoding patched gene
 other than Drosophila patched gene, useful for preparing transgenic
 laboratory animals and to knock out patched protein in embryonic stem
 cells.

XX PT

XX Disclosure; Page 8-10; 40pp; English.

XX PS

The invention relates to a DNA sequence other than present in chromosome
 encoding a patched (ptc) gene other than Drosophila patched gene. A cell
 expressing the DNA sequence is useful for producing patched protein, by
 growing the cell expressing the DNA sequence, where the patched protein
 is expressed and isolating the patched protein free of other proteins.
 The cell expressing the DNA sequence is also useful for screening

XX CC

XX CC

XX CC

XX CC

XX CC

XX CC

candidate compounds for binding affinity to the patched protein, by
 combining the candidate protein with the cell expressing the DNA sequence
 where the DNA sequence comprises the entire coding sequence under the
 transcriptional regulation of the transcriptional initiation region and a
 transcriptional termination region functional in the cell, expressing the
 patched protein in the cell and assaying for the binding of the candidate
 compound to the patched protein. The above cell is useful for screening
 candidate compounds for agonist activity with the patched protein. The
 DNA sequence is useful for producing all or portions of the patched
 protein, as probes for research, diagnosis, binding of hedgehog protein
 for its isolation and purification and in gene therapy. The DNA sequence
 is also useful as primers for investigating other species and for
 isolating genes from various mammalian sources of interest, particularly
 from humans or from domestic animals. The DNA sequence is further useful
 for preparing transgenic laboratory animals and to knock out the PTC
 protein in the embryonic stem cells, so as to produce hosts with single
 functional patched gene. The present sequence represents the amino acid
 sequence of the buckeye patched gene PTC product

XX Sequence 1311 AA;

Query Match 7.6%; Score 526.5; DB 7; Length 1311;

Best Local Similarity 22.2%; Pred. No. 1.1e-38;

Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LIGQFFQGGTWTWASWPLTILVLSIPVVALAAGLVFTLTTPDVELMSAPNSQARSEKA 395

Db 72 ILGCFLOQDAGKVL--VAILVLSFCV-----GLKSAQIHTVDQMWQEGREAELEK 124

QY 396 FHDQHFQFP--FRTNQVILTAPNRSYRYDSLLGPNFSGILDLDLLELLEQLERLHL 454

Db 125 YTAQALGADSSHTOLVI---QTAKDPDVSLLHP---GAL-----LEHL 162

QY 455 QVMSPEAQORNI-----SLQDICVAPLNPD-----NTSLVD---CCINSILQVQNN 497

Db 163 KVVHAARTVTVMYDIEWRLKDCYSPSIDPEGVHHIESIDNVIPCAITTPLOCFWEG 222

QY 498 RTLL-----LLTANQTLMGQTSQ-----515

Db 223 SKLLGPDYPIYVPHLKHKLQWHLNPLFVVEVVKLKQFPPLSTIEAYMKRAGITSAYNK 282

QY 516 ---VDWKDHFLYC--ANAPLTFQDGTALALSCMADYGAPVFPF-----LAIG-----558

Db 283 KPCLDPTDP--HCPATAP--NKKSGHIPDVAEELSHGCGYGAAYAMHWPQLIVGGATRNS 339

QY 559 ---YKGDYSEAEALIMTFLSNYPAGDPRLAQAKLWEAELEEMRAFORWAGFPQT 614

Db 340 TSALEKARXLQTVQLMGEREMYEWADHYKVHQIWNQOEKAAAVLDAMQKFAAEVRKI 399

QY 615 FT-----AERSLEDEINRTTAEDLPIPATSYIVIFLYISALGSYSSSRVM 661

Db 400 TTSGSVSSAYSFPFSTSTLNDILGKPFSEVSLKNILGVMFMLIYAVTL---IQW-RDP 455

QY 662 VDSKATLGLGVAVVVLGAVMAAMGPFYSILGRSSILVILQVPPVLVLSVGADNIFIVLEY 721

Db 456 IRSQAGVGIAGVLLLSITVAAGLGFPCALLGPFNASTQIVPFLALGLGVQDMFLITHTY 515

QY 722 QRLPRPGEPREHVHIGRALGRVAPSMALCSLSEACFFLGALTMPAVRTFALTSLAVI 781

Db 516 ---VEQAGVDPREBRTGLVLKKSGLSVLASLGNVMAFLAAALLPIPAFRVFLQAAAILLL 573

QY 782 LDPLQLMSAFVALLSLDSKREASRLDVCCCKVKPQELPP-----820

Db 574 FNLSGILVFPNMTISLDLRRRSARADLLCCLMPESPLPKKKI PERAKTRKNDKTHRIDT 633

QY 821 -----PQGGEL-----LIGFFQKAVAPAPLHLLITRGVILLFLFVGSLSYNC 865

Db 634 TRQLDPOVSENVVTTCCLSVSLTKWAKNQVAFPIRMRAVKVTSMLIALIIVLTSSVWGAT 693

QY 866 HISVLQDELALPKDSYLLDYFLELNR---VEEVCAPVYFVTLGLGVNSESSEAGNAICSS 922

Db 694 KYKQGLDLDIVPENT---DEHFFLSRQEKYFGF-----YN-----MYAVTQ-- 732

QY	923	AGCNHFF--TQKIQYATEPPEQSYLAIP-----ASSWVDDFIDWL-----	961
Db	733	---GNFEPYTNQKLLY--EYHQ-Q-FVRIPIIKNDNGGLTKFWSLFRDWLLDLQVAFDK	786
QY	962	-TPSSCCRLYISGNKQKFCPSTVNSLNCNCHMSITWGSVRPSVEQ-----	1007
Db	787	EVASGCI-----TQEWYCKNASDE-GILAYKLMVQTHVDNPDIKSLITAGHRLVDKD	838
QY	1008	-----PHKYL-PWFLNDREPNKCPKGGI-----AAYSTSVNLTSDGQVLASRFM	1050
Db	839	GIINPKAFYNYLSAWATNDALAYGASQGNLKPQQRWTHSPEDVHLEIKKSSPLIYQLP	898
QY	1051	AHKFKLNSQDYTEALRAARELANITADLRKKVPGTDPAFVFPYTTINVFYEQYLILP	1110
Db	899	FYLSGLSDTXXSKTLIRSVRDLCIKEYA--KGLPN-----FPGSIPFLFWEQYLYLRT	949
QY	1111	EGFLMILSLVPTFAVSCLLGLDLRLSGLLNLSIVMLIVDTVGFMALWDISYNAVSLIN	1170
Db	950	SLLALACALAAVF-IAVWVLLNNAVAVTLATLVTLQILGWALLGVKLSAMPVAVL	1008
QY	1171	LVSAGVMSVEFVSHITRGSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLA	1230
Db	1009	LVLALGRGVHTVHLCLGFVTSIGC-----KRRASLALES-VLAPV-----VHGALAAALA	1059
QY	1231	KAQL-----IOIFFFRNLNLIITLGLLHGLVFLPVILSVYGP--DYNPALALEQKEAE	1281
Db	1060	ASMLAAASECGFVARLFLRLLLDIVFLGLTDLGLLFFPILVLSILGPAAEVRPI-----	1110
QY	1282	EAVAAMVWASCNPSPSRVST	1301
Db	1111	-----EHPERLST	1118
RESULT 67			
ID	AD94209	AD94209 standard; protein; 1311 AA.	
XX	AC	AD94209;	
XX	DT	12-FEB-2004 (first entry)	
XX	DE	Peacock butterfly ptc polypeptide.	
XX	KW	Peacock butterfly; patched gene; ptc; developmental abnormality;	
XX	KW	proliferative disorder; tumour; oncogenic patched mutation;	
XX	KW	patched-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma;	
XX	KW	meningioma; fibroma; cancer.	
XX	OS	Junonia coenia.	
XX	PN	US6551782-B1.	
XX	PD	22-APR-2003.	
XX	PF	28-NOV-2000; 2000US-00724631.	
XX	PR	07-OCT-1994; 94US-00319745.	
XX	PR	06-OCT-1995; 95US-00540406.	
XX	PR	31-MAY-1996; 96US-00656055.	
XX	PR	22-AUG-1997; 97US-00918658.	
XX	PA	(STRD) UNIV STANFORD.	
XX	PA	(REGC) UNIV CALIFORNIA.	
XX	PI	Scott MP, Goodrich LV, Johnson RL, Epstein E;	
XX	DR	WPI; 2003-615308/58.	
XX	DR	N-PSDB; AD94208.	
XX	PT	Assay for phenotyping patched status of cell in mammalian (preferably	
XX	PT	human) cell sample, involves detecting presence or absence of aberrant	
XX	PT	modification or mutation of a patched gene, and mis-expression of patched	
XX	PT	gene.	
XX	PS	disclosure; SEQ ID NO 4; 57pp; English.	
XX	CC	The invention relates to an assay for phenotyping the patched status of a	
XX	CC	cell, involving detecting in vitro, in a sample of mammalian cells, the	
XX	CC	presence or absence of a genetic lesion in a patched gene (ptc)	
XX	CC	characterised by at least one of an aberrant modification or mutation of	
XX	CC	a patched gene or mis-expression of the patched gene. The invention also	
XX	CC	relates to a method for diagnosing a genetic predisposition of an animal	
XX	CC	to a developmental abnormality or a proliferative disorder, where the	
XX	CC	developmental abnormality or proliferative disorder correlates with	
XX	CC	aberrant expression or activity of a patched gene or gene product,	
XX	CC	involving detecting in vitro the presence of a predisposing mutation in a	
XX	CC	patched gene in cells of the animal. Characterising the phenotype of a	
XX	CC	tumour, involves detecting the presence of an oncogenic patched mutation	
XX	CC	in cells of the tumour, where the presence of the oncogenic mutation	
XX	CC	indicates that the tumour has a patched associated phenotype. The assay	
XX	CC	is useful for phenotyping the patched status of the cell in a mammalian	
XX	CC	cell sample obtained from a human patient. The disorders that can be	
XX	CC	detected include basal cell nevus syndrome (BCNS), carcinomas,	
XX	CC	meningiomas and fibromas. This sequence represents a butterfly ptc	
XX	CC	polypeptide of the invention.	
XX	SQ	Sequence 1311 AA;	
Query Match 7.6%; Score 526.5; DB 7; Length 1311;			
Best Local Similarity 22.2%; Pred. No. 1.1e-38;			
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;			
QY	336	LLGQFFQCGWTWASWPLTILVSVIPVALAGLVTELTTPVELWSAPNSQARSEKA	395
Db	72	ILGCFLOGDAGKVLFF--VAILVLSTFCV-----GLKSAQIHTRVDQLWVQGGRLAEK	124
QY	396	PHDQHFQFP-FRTNOVILTAPNRSRYDSLLGPKNFGSILDLLELLELLEQLRHL	454
Db	125	YTAQALGEADSTHQLVI---QTKAPDVSLLHP-----GAL-----LEHL	162
QY	455	QVWSPEAQRI-----SLQDICYAPLNPD-----NTSLYD-----CCINSLOYQNN	497
Db	163	KVHAATRVTHMYDIEWRLKDLCSFSPIDFEGYHIESIIDNVIPCAITPLDCFWE	222
QY	498	RTLL-----LTTANQTLMGOTSO-----	515
Db	223	SKLLAGPDYPIYVPHLKHKLWHLNPLEWVEEVKGLKFPPLSTIEATMKRAGITSYMK	282
QY	516	---VDWKDHFYLC-ANAPLTPKDGKTALALSCMADYAGAPVFPF-----LAIG-	558
Db	283	KPCLDPTD--HCPATAP-NKSGHIPDVAELSHGCGYFAAAYMHWPEQLIVGATRNS	339
QY	559	---YKGRDYSEAEALIMTFSLNYPAGDPRLAQKLWEBAFLBEMRAFORMAGMFQVT	614
Db	340	TSALRKARXLQTVQLMGEREMYEWADHYKHQIGWNQEKAAAVALDAWRKFAAEVRKI	399
QY	615	FT-----AERSLEDEINRTTAEDLPFATSVIVIFLYISALGSYSSWSRV	661
Db	400	TTSGSVSAYSFPYFSTSTLNDILGKSEVLSKNILGYMFMILIVAVTL---IQW-RDP	455
QY	662	VDSKATLGLGVAVVVLGAVMAAGFFSYLIGIRSSILVILQWVPFLVLSVGADNIFVLEY	721
Db	456	IRSQAGVGIAGVLLLSITVAAGLGFALLGIPFNASSQIVPFLALGLGVQDMFLTHTY	515
QY	722	QRLPRRPGEPREVHIGRALGRVAPSNMLCSLSEATCFPLGALTMPAVRTALTSLGLAVI	781
Db	516	--VEQAGDVPREERTGLVKKSGLSVLLASLCNVNAFLAAALLPIPAFRVFLQQAILL	573
QY	782	LDFLLOMSAFVALLSLDSKQEAESLDVCCCKVPQOELPP-----	820
Db	574	FNLGSILLVFPAMISLDLRRRSAAADLLCCLMPESPFLPKKXI PERAKTRKNDKTHRIDT	633
QY	821	-----PGOGEGL-----LLGFFQKAYAPFLLHMTITRGVVLILLFALFGVLSYMC	865
Db	634	TROPDPDVSENVTKTCCLSVSLTKWAKNQAFPIMRPAVKVTSMLALIAVILTSVWGAT	693

QY 866 HISVGLDQELALPKDSYLLDYFLPLNR---YFEVGAPVYFVTVTLTGNFSSEAGMNAICSS 922
 Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYFGF-----YN-----MYAVTQ- 732
 QY 923 AGCNNSF--TKIQIYATEPPEQSILAIP-----ASSWVDPIDWL----- 961
 Db 733 ---GNFEYPTNQKLLY--EYHDQ-FVRIPIIKNNGGLTKFWSLFRDWLLDLQVAFDK 786
 QY 962 -TPSSCCRLYISGPNKDKPCPSTVNSLNCIKNCMSITMGSVRPSVBQ----- 1007
 Db 787 EVASGCI-----TQBYCKNASDB-GILAYKLMVQTGHVDNPIDKSLITAGHRLVDXD 838
 QY 1008 -----FKYLL-PWFLNDRPNIKCPKGGI-----AAYSTVNLSDDGQVILASRFM 1050
 Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPOQRWIHSPEVDVHLEIKKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITIVFEYQVILTILP 1110
 Db 899 FYLSGLSDTXSIKTLIRSRVDLCLKYEA--KGLPN-----PPSGIPFLFWQYILYLT 949
 QY 1111 EGLFMLSCLIVPTFAVSCLLGLDLRSGLNLISIVMILVDTVGFMAWDISYNAVSLIN 1170
 Db 950 SULLALACALAAVF-TAVMVLLNNAVAVLTLALATLVLLQGLVWALLGVKLSAMPAVL 1008
 QY 1171 LVSAGMSVEFVSHITRSFAISTKPTWLERAKETISMGSAVPAVAMTNLPGILVGLA 1230
 Db 1009 LVLAIGRGVHFTVHLCLGFTVSGC---KRRRASIALES-VLAPV---VHGALAAALA 1059
 QY 1231 KAQL-----TQIFFRNLITLGLLHGLVPLVPLVILSYVCP--DVNPALALEOKRAE 1281
 Db 1060 ASMLAASECGFVARLFLRLDLLVFLGLIDGLLFFPPIVLSILGPAAEVRPI----- 1110
 QY 1282 EAAVAVMVASCPNHPSRVST 1301
 Db 1111 -----EXPERLST 1118

RESULT 68
 ADH62716
 ID ADH62716 standard; protein; 1311 AA.
 XX AC ADH62716;
 DT 25-MAR-2004 (first entry)
 XX DE Butterfly patched (bptc) homologue protein.
 XX KW PTC; patched; embryonic development; cellular regulation;
 XX KW signal transduction; butterfly.
 XX OS Junonia coenia.
 XX FT Key Location/Qualifiers
 XX FT Misc-difference 348 /label= Unknown
 XX FT Misc-difference 908 /label= Unknown
 XX FT
 XX PN US2003148388-A1.
 XX XX
 XX PD 07-AUG-2003.
 XX XX
 XX PP 03-JAN-2001; 2001US-00754032.
 XX XX
 XX PR 07-OCT-1994; 94US-00319745.
 XX PR 06-OCT-1995; 95US-00540406.
 XX PR 20-OCT-1997; 97US-00954668.
 XX XX
 XX PA (STRD) UNIV STANFORD.
 XX XX
 XX PI Scott MP, Goodrich LV, Johnson RL;
 XX XX
 XX DR WPI; 2003-897566/82.

XX

PT

PT

PT

PT

XX

XX

XX

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Query Match

Best Local Similarity

Matches

336

72

396

125

455

163

498

223

516

283

559

340

615

400

662

456

722

516

782

574

821

634

866

694

923

733

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

New DNA sequences encoding patched protein from organisms including human and mouse are used to elucidate embryonic development and cellular regulation associated with patched signal transduction and identify patched ligands.

Disclosure; SEQ ID NO 4; 40pp; English.

The present invention relates to a DNA sequence encoding a patched (PTC) gene. The invention is used to elucidate embryonic development, cellular regulation associated with signal transduction by the patched gene, identify agonists and antagonists of that signal transduction, identify and isolate ligands of patched protein and assay for levels of PTC transcription of the patched gene. The present sequence is butterfly patched protein.

Sequence 1311 AA;

Query Match 7.6%; Score 526.5; DB 7; Length 1311;
 Best Local Similarity 22.2%; Pred. No. 1.1e-38;
 Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 ILGQFQGGTGWASWPLTILVLSVTPVALAAGLVFTLTPDVELWAPNSQARSEKA 395

Db 72 ILGCFLQGDAGKVLV--VAILVLSTFCV-----GLKSAQIHTRVQDLVQEGGRLEALK 124

QY 396 FHDQHFQPP-FRTNOVILTAPNRSSVRYDSLLGPKNPSGILDLDLLELLELQERLHL 454

Db 125 YTAQALGEADSTHQDLV---QTAKDPDVSLLHP-----GAL-----LEHL 162

QY 455 QWMSPEAQRNI-----SLQDICYAPLNPD-----NTSLYD-----CCINSILQYFQNN 497

Db 163 KVVHAATRVTVHMYDIEMWELKDCYSPSPDPGEYHHIESIIDNVIPCALITPLDCFWEG 222

QY 498 RTLL-----LLTANOTLMQTSQ--- 515

Db 223 SKLLGPDYPIYVPHLKHKLQWTHLNPFLVYVEVKLKFQPLSTIEAYMKRAGITSAYMK 282

QY 516 ---VDMKDHFLYC-ANAPLTPKDGITALALSCWADYCAPVPPF-----LAIGG----- 558

Db 283 KFCPLDTPD--HCPATAP-NKSGHIPDVAAELSHGCGYGAAYMFWPQOLIVGGATRNS 339

QY 559 ---YKQDYSEAEALIMTFSLNYPAGDPRLQAQKLWEEAFLEENRAFQRMAGMFQVT 614

Db 340 TSALRKARXLQTVVQLMGEREMVEYWDHYKVHQIGMNOEKAAAVLDAMQKFAAEVRKI 399

QY 615 FT-----AERSLEDEINKTAEADLPFATSYIVIFLYISLALGSSVSSSRVM 661

Db 400 TTSGSVSSAYSFPFSTSTNDILGKFSEVSLKNIILGYMFMFLIYVAVTL---IQW-RDP 455

QY 662 VDSKATLGLGGVAVVLGAVMAAGPSPYLGIRSSLVILQVPPFLVLSVGNADNIFIVLEY 721

Db 456 IRSQAGVGIAGVLLSITVAAGLGCALLGIPFNASSTQIVPPLAALGLGQDMFLTHTY 515

QY 722 QRLPRRPGEPREHVHIGRALGRVAPSMILCSLSEALICFFLGCALTMPMPAVTFALTSLGLAVI 781

Db 516 ---VEQAGDVPREERTGLVKKSGLSVLASLNCWMAFLAAALLPPIPAFVFCIQAAILLL 573

QY 782 LDFLLQMSAFVALLSLDSKRBASRLDVCCCKVQBELPP----- 820

Db 574 FNLGSILLVFPAMISLDLRRRSARADLLCCLMPSPPLPKKKIPERAKTRKNDKTHRIDT 633

QY 821 -----PGQEGE-----LLGFFQKAYAPFLHWTIRGVTLLFLALFVSLYSMC 865

Db 634 TRQPLDPDVSENVTKTCCLSVSLTKWAKNQYAPFIMRPAPVKVTSMLALIAVILTSVWGAT 693

QY 866 HISVGLDQELALPKDSYLLDYFLPLNR---YFEVGAPVYFVTVTLTGNFSSEAGMNAICSS 922

Db 694 KYVDGLDLTDIVPENT---DEHEFLSRQEKYFGF-----YN-----MYAVTQ- 732

QY 923 AGCNNSF--TKIQIYATEPPEQSILAIP-----ASSWVDPIDWL----- 961

Db 733 ---GNFEYPTNQKLLY--EYHDQ-FVRIPIIKNNGGLTKFWSLFRDWLLDLQVAFDK 786

XX The invention describes an assay for phenotyping the patched status of a
CC cell comprising detecting in a sample of mammalian cells the presence or
CC absence of a genetic lesion having aberrant modification or mutation of a
CC patched gene or mis-expression of the patched gene. The assay is useful
CC for diagnosing a genetic predisposition of an animal, e.g. basal cell
CC nevus syndrome, predisposition for developing tumour, i.e. carcinoma,
CC meningioma, medulloma, or fibroma. A genetic construct encoding a patched
CC polypeptide is used to treat an animal having a disorder comprising loss
CC of function of a wild-type patched gene, such as cancer, and can enhance
CC patch function in e.g. wound healing and aging. This sequence is encoded
CC by the butterfly patched gene.

XX SQ Sequence 1311 AA;
Query Match 7.6%; Score 526.5; DB 8; Length 1311;
Best Local Similarity 22.2%; Pred. No. 1.1e-38; Indels 307; Gaps 46;
Matches 258; Conservative 175; Mismatches 420;

QY 336 LIQGFQGGTWTASWPLTILVSVVVALAAGLVFTELTDPVELWSAPNSQARSEKA 395
DB 72 ILGCFQGGAGKVLFP--VAILVLSSTFCV-----GLKSAQIHTRVDQLMVQEGGRLEALK 124
QY 396 FHDQHFQFP--FRINQVILTAAPNSRYSDLSLLGPKNFSGILDLDLLELLELQERLHL 454
DB 125 YTAQALGEADSSHTQLVI---QTAKDPDVSLLHP---GAL-----LEHL 162
QY 455 QVMSPEAQRNI-----SLQDICYAPLNDP-----NTSLYD---CCINSLLQYFQNN 497
DB 163 KVHAATRVTVHMYDIEWRLKDLCSYSPIDPFEGYHHIESIINDNVIPCAITLDCFWEG 222
QY 498 RTLL-----LLTANQITMGQTSQ--- 515
DB 223 SKLLGPDYPIYVPHLKHKLQWHLNPLEVVEVKLKFQPLSTIEAVMKRAGITSAYMK 282
QY 516 ---VDWKDHFLYC--ANAPLTPKQGTALALSCMADYDAPVFPF-----LAIGG----- 558
DB 283 KPCLDPTDP--HCPATAP--NKSGHIDPVAEAELSHGCGYFAAAAYMHWPQELIVGATGNS 339
QY 559 ---YKDYSEAEALIMTFSLNYPAGDPLAOKLWEAEFLERMAFORRMAGMFOVT 614
DB 340 TSALRKARXLQTVVQLMGEREMYEYADHYHQHIGNOEKAAAVLDQWQKFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISIALSGYSWSRNM 661
DB 400 TISGSVSSAYSFPFSTSLNDILGKFSEVSLKNILGYMFMLIXVAVTL---IQW-RDP 455
QY 662 VDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSVSGADNIFIVLEY 721
DB 456 IRSQAGVGIAGVLLLSITVAAGLGCALLGIPFNASTSTQIVPFLGALGVQDMFLTHTY 515
QY 722 QRLPRPGEPREVHIGRALGVAPSMGLCSLSEACFFLGLALTMPAVRTFALTSLGLAVI 781
DB 516 --VEQAGDVPREERTGLVLKXSGLSVLASLGNVWAFLLAALLPFPFVFLQRAILL 573
QY 782 LDFLLQMSAFVALLSLDSKROEASRLDVCCVKKPQELPP----- 820
DB 574 FNLGSILLVFPAMISLDLRRSARADLLCCLMPESPFPKKKIPERAKTRKNDKTHRIDT 633
QY 821 -----PGQGEGL-----LLGFFQKAYAPFLLHWTITRGVLLFLFALFVLSYSC 865
DB 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALFVILTSVWGAT 693
QY 866 HISVGLDQELALPKDSYLLDDYFLNRR---YFEVGAFFVYVTTTLGYNFSSAGMNAICSS 922
DB 694 KYKQGLDLTDIVPENT---DEHFLSRQEKYFGF-----YN-----MYAVTQ- 732
QY 923 AGCNFESF--TKIQYATFEPQSYLAIP-----ASSWDDFIWL----- 961
DB 733 ---GNEFYTNKLLY--BYHDQ--FVRIENIKNDNGGLTKFMLSIFRDLWLLDQVAFDK 786
QY 962 -TPSSCCRLYISGPNKDKFCPTVNSLNCNCKMSITMGSVRPSVEQ----- 1007

QY 962 -TPSSCCRLYISGPNKDKFCPTVNSLNCNCKMSITMGSVRPSVEQ----- 1007
DB 787 EVASGCI-----TQEWCKNASDE--GILAYKLVQVGHVDNPIDKSLITAGHRLVDKD 838
QY 1008 -----FHXYL--PWFLNDRENIKPKGGL-----AYSTSVNLTSDGGVLSARFM 1050
DB 839 GIINPKAFYNLSAWATNDALAYGASQGNLKPQQRWIHSPEDVHLEIKKSSPLIYTQLP 898
QY 1051 AYHKPLKNSQDYTEALRAARELAANIADLRKVPCTDPAFEVFPVTTITNVFYEQYLTLP 1110
DB 899 FYLSGLSDTSIKTLIRSVRDLCKLYEA--KGLFN-----FPSGIPFLFWEQYLYLRT 949
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSRLSGLNLLSLVMILVDTVGFMALMDISYNAVSLIN 1170
DB 950 SILLALACALAAVF--IAVWVLLNNAVAVTLATLVLQGLGWMALLGVKLSAMPV 1008
QY 1171 LVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLA 1230
DB 1009 LVLAIGRGVHFTVHLCGLGFTVSIGC---KRRASLALES-VLAPV-----VHGALAAALA 1059
QY 1231 KAQL-----IQIFFRLNLLITLGLLHGLVLPVILSVYGP--DVNPALALEQKRAE 1281
DB 1060 ASMLAASECGFARLFLRLLDIVFLGLTDGLLFFPIVLSLIGPAAEVRPI----- 1110
QY 1282 EAVAAVMVASCPNHPSRVST 1301
DB 1111 -----EHPERLST 1118
RESULT 69
ADE48974
ID ADE48974 standard; protein; 1311 AA.
XX ADE48974;
AC ADE48974;
XX ADE48974;
DT 29-JAN-2004 (first entry)
XX Butterfly patched protein (ptc).
DE cytostatic; vulnary; gene therapy; phenotyping; patched status;
KW patched gene; genetic predisposition; basal cell nevus syndrome; tumour;
KW carcinoma; meningioma; medulloma; fibroma; cancer; wound healing; aging;
KW butterfly; patched gene; ptc.
XX Junonia coenia.
OS
XX US2003186309-A1.
XX
XX 02-OCT-2003.
XX 22-APR-2003; 2003US-00421446.
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX 22-AUG-1997; 97US-00918658.
XX 28-NOV-2000; 2000US-00724631.
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX (REGC) UNIV CALIFORNIA.
XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX WPI; 2004-041193/04.
XX N-PSDB; ADE48973.
XX Phenotyping the patched status of a cell for diagnosing a genetic
XX predisposition for a tumor comprises detecting the presence or absence of
XX a genetic lesion having aberrant modification, mutation or mis-expression
XX of the patched gene.
XX Disclosure; SEQ ID NO 4; 60pp; English.
PS

Db 787 EVASGCI-----TQEVCKNASDE-GILAYKLMVQTGHVNDPDKSLITAGHRLVDKD 838
 QY 1008 -----FKYL-PWFLNDRPNKCPKGG-----AAYSTSVNLTSDQGVLASRFM 1050
 Db 839 GIINPKAFYNYLSAWATNDALAYGASQGNLKPQQRWIHSPEVDVHLEIKKSSPLIYTQLP 898
 QY 1051 AYHKPLKNSQDTEALRAARELANITADLRKVGTDPDFVFPYTTITNVFEQVLTILP 1110
 Db 899 FYLSGLSDTXSKTILRSVRDCLKEYA--KGLPN-----FPGSIPFLFEQVLYLRT 949
 QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLNLLSVIMILVDTVGFMAWMDISYNAVSLIN 1170
 Db 950 SILLALACALAAVF-IAVWVLLNAAVAVLVITATLVQLGVALGKLSAMPVL 1008
 QY 1171 LVSAGVMSYFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLA 1230
 Db 1009 LVLAIGRGVHTVHLCLGFPVTSIGC-----KRRASLALGS-VLAPV-----VHGALAALA 1059
 QY 1231 QAOL-----IQIFFRLNLIITLGLLHGLVFLPVILSYVGP--DYNPALALSKQRAE 1281
 Db 1060 ASMLAASECGFVARLFLRLLDIVFLGLTDLGLLFFPVLVLSILGPAAEVRPI----- 1110
 QY 1282 EAVAAVMVASCNHPDSRVST 1301
 Db 1111 -----EHPERLST 1118

RESULT 70

ADN22811
 ID ADN22811 standard; protein; 1405 AA.

AC ADN22811;

DT 02-DEC-2004; (first entry)

DE Bacterial polypeptide #5464.

Recombinant DNA construct; transformed plant; improved plant property;
 cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 pathogen tolerance; pest tolerance; plant disease resistance;
 cell cycle pathway modification; plant growth regulator;
 homologous recombination; seed oil yield; protein yield; carbohydrate;
 nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 bacterial polypeptide.

OS Bacteria.

US2003233675-A1.

18-DEC-2003.

20-FEB-2003; 2003US-00369493.

21-FEB-2002; 2002US-0360039P.

(CAOY/) CAO Y.

(HINK/) HINKLE G J.

(SLAT/) SLATER S C.

(CHEN/) CHEN X.

(GOLD/) GOLDMAN B S.

Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

WPI; 2004-061375/06.

New recombinant DNA construct comprising a promoter positioned to provide
 for expression of a polynucleotide encoding a polypeptide from a
 microbial source, useful for producing plants with improved properties.

Claim 1; SEQ ID NO 5464; 122pp; English.

The invention relates to a recombinant DNA construct comprising a
 promoter functional in a plant cell, where the promoter is positioned to

CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 1405 AA;

Query Match 7.4%; Score 514; DB 8; Length 1405;

Best Local Similarity 20.2%; Pred. No. 1.8e-37;

Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSFTHLLGQPFQCGWTV--ASWPLTILVLSVIVVVAALAGVLTFTLTD 378

Db 100 RKGATGNRYALYSRSLIOKLLFALGNTVHRNAWSI-ILAVSMIFAVC-CYGLQYVHIETD 157

QY 379 PVELMSA-----PNSQARSEKAFHQHGFPP-RTN-----QVILTAPNRSS 419

Db 158 IVKLWVAOGGRLDDELNPLNKEAMRNVTGDS--GPBLPRENGLGCGVQLIQPEYEG 215

QY 420 YRYDLSLLGPNFSGILDLDLLELLELQERLHLQVMSPEAQRNI-----SLQDICY 472

Db 216 --QDALAAGP-----LLKHVEIMKHIAFNVSVHGVMDSLDICF 253

QY 473 APLNDP-----NTSLYD-----CC-INSLLQYFQNNRTU-----LLL 503

Db 254 KPAPPSVAADSAASLSGDVIDKIVPCITWITPIDCFWEGSKALGPHPSLPKSSILGLMGL 313

QY 504 TANQTLMGTSQVDMKD-----HFLYCANAPLTP----- 532

Db 314 SS-----LSDGMIRWSDFDPIAVIDEIHRSFNLGSHYTFEFERAGVSHGMDRCPIDPLD 369

QY 533 -----KDGITALALSCMADYCAPVPFPFLAIGGYKGYSEAR 568

Db 370 ECPPMKNYFDVCPHIDRVREIAKYGYTELEBEKKKSGYSFFDL---GRKKREAGDOP 426

QY 569 ALI-----MTFSLNYP-----AGDP 584

Db 427 KMIHPAQPADSIPTIEDAVPAQVPVSTAPITPTTTLSPEARAABEKKQKARELKDYC 486

QY 572 ----- 546

Db 487 KSYRKSAPFWLKKNDKWPVMSNNYFQNVDYAAEMTGGCGSPASVNLNWPEDMILGNP 546

QY 585 RLQAQ--AKL-----WEAFLEB-MRA 602

Db 547 RRKKGKLSGNALQSVPLVASPADVFLRFKQKPGNRSMKTGLDMDAWNETAQAQVLOA 606

QY 603 FQRMAGMFQVTFTAERSL-----EDEINRTTAEDLPAPTS-----YIV 642

Db 607 WQR-----NFT--KSLYNHKAHVDDGNERRTLH--PLASTSIADMLBEFCQFNVTI 654

QY 643 IF-----LYISLAGSYSSWSRVWVDSKATLG--LGGVAVVLGAVMAAGFPYSYLGRSL 696

Db 655 ILAGYALMLAYAVITQARFDNCLPATESSMGLALAGVLVITFASVAGLGLATWFGIFRNA 714

Db 400 TTSGSVSSAYSFPSTSTLNDILGKPSSEVSLKNIILGYMFMIIYVAVTL-----IQW-RDP 455
QY 662 VDSKATLGLGGVAVVLGVAAMAGFSGYLGIRSSILVILQVWPFVLVSVGADNIFIVLEY 721
Db 456 IRSAGVGLVGLLSITVAAGLGFALLGIPNASSSTQIVFPLALGLGVQDMFLTHY 515
QY 722 QBLPRPGEPREVHIGRALGRVAPSMLLCSLSAICFFLGLATPMPAVNTFALTSGLAVI 781
Db 516 --VEGAGDVPREERTGLVKSGSLSVLLASLNCVMAFLAAALLPAPFRVFCLOAAILL 573
QY 782 LQFLQMSAFVALLSLDSKQASRLDVCCCKVQPELPP----- 820
Db 574 FNLGSLVFPAMISLDLRRRGAAPADLCCLPSPLEPKKIPERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGEL-----LLGFFKAYAPFLLHWITRGVILLFLALFGLSVLSYMC 865
Db 634 TRQPLDPDVSENVTKCCLSVSLTKWAKNQYAFIMRPAVKYTSMLALIAVILTSWGAT 693
QY 866 HISVLGDQELALPKDSYLLDYFLNRR-----YFEVGAPVYFVTTLGVNFSSEAGMAICSS 922
Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYRGF-----YN-----MVAVTQ- 732
QY 923 ACNNFSP--TKIQYATEFPQSYLAIP-----ASSWDDFDWL----- 961
Db 733 ---GNFEYPTNQKLLY--EYHQDQ-FVRIPNIIKNDNGGLTKFMLSFRDMLDLQVAFDK 786
QY 962 -TPSSCCRLYISGPNKDKPCSTVNSLNCMKMSITMGSVRPSVEQ----- 1007
Db 787 EVASGCI-----TOEYWCXNASDE-GILAYKLVQVGHVNDPNIDKSLITAGHRLVDKD 838
QY 1008 -----FKHYL-PWFLNDRPNIKCPKGL-----AAYSTSVNLTSDQGVLASRPM 1050
Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQORWIHSPEDVHLEIKKSSPLIYVTLQ 898
QY 1051 AYHKELKNSQDTEALRAARELANITADLRKVGTDPAFYVPPYTIITVFEQVLTILP 1110
Db 899 FYLSGLSDTDSIKTIRSVRDIJCLYEA--KGLPN-----FPGSIPFLFEQYL-YLR 948
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSRGLNLSITVMILVDTVGFPMALWDISYNAVSLIN 1170
Db 949 TSLLALACALGAVFIAMVLLNNAWAVLVTLATLVQLLGVALLGVKLSAMPVYL 1008
QY 1171 LVSAVGMSEVESHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLA 1230
Db 1009 LVLAIGRGVHFTVHLCLGFPVTSIGC---KRRASLALLES-VLAPV---VHGALAAMLA 1059
QY 1231 KAQL-----IQIPFRNLNLTLLGLLHGLVFLPVLISYVGP--DVNPALALSQKRAE 1281
Db 1060 ASMLAASEFGFVARLFLRLALLALVFLGLTDGLLFFPVLISILGPAAEVRPI----- 1110
QY 1282 EAVAAMVAVASCPNHPRSVST 1301
Db 1111 -----EHPERLST 1118
RESULT 72
ABB59092
ID ABB59092 standard; protein; 1286 AA.
XX AC ABB59092;
XX DT 26-MAR-2002 (first entry)
XX DE Drosophila melanogaster polypeptide SEQ ID NO 4068.
XX KW Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical.
XX OS Drosophila melanogaster.
XX PN WO200171042-A2.
XX XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.
XX PF
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
XX N-PSDB; ABL03195.
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX Disclosure; SEQ ID NO 4068; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1286 AA;
Query Match 6.9%; Score 474; DB 4; Length 1286;
Best Local Similarity 20.4%; Pred. No. 8.5e-34;
Matches 245; Conservative 177; Mismatches 457; Indels 322; Gaps 42;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSVIPVVALAAGLVTELTDPV 380
Db 43 KKGARSGRTAIVLRVFPQSHLETGSSVQKHAGKVLFAVLVLTFCVGLKSAQIHSKVH 102
QY 381 ELWAPNQARSEKAFPHQHFQ-PFRTNQVILTA---PNRSYRYDSLLGLPKNPSGIL 436
Db 103 QLWIOEGGRLEABLAYTKTIGEDSATHQLLIOTTHIDPNASVLHPQALL-----A 153
QY 437 DLIDALLELEQLERLHQLVMSPEAQRNLSLDICVAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKTATAVKHYLDTE-W-----GLRDMCMFSTPSPFEGIYYIEQILRHLPICS 204
QY 487 INSLLOQYFQN-----NRTLILLLTAN-----QTLMGQTSQVNDKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVIPGLNQRLWLTILNPASVMQYKQKMSSEKISPDFTV 264
QY 522 FLYCANAPLTFKDGTA-LALSQM-----ADYGAPVFPFPAIG--GY 559
Db 265 EGYMKRAAI-----GSGYMEKPCINPLNPNCPDAPFNKNSQTPDPVGA-----ILSGCYGY 316
QY 560 KGDYSEAEALTMTPSLANNYPAGDPLAQA----- 589
Db 317 AAKHMHVPEELIVGAKRN-RSGHLRKAQALQSVOQLMTEKEMQMDQONKYVHHLGWTO 375
QY 590 -----KLWBEAF---LEEMRAFQRRMAGMFQVTFTAERSLEDEINRTTARDLPFATS 639
Db 376 EKAAEVLNMQNRFSEVEQLLRKQSRATNVDYIVFPSSAALDDILAKFSHPSALSIVIG 435
QY 640 YIVIFLYTSLALGSYSWSRVNVDKATLGLGVAVLGVANVAMGFFSYLGIIRSLVL 699
Db 436 VAVTVLY---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAAGLSALLGIVFNAAST 491
QY 700 QVVPFLVLSVGADNIFIVLEVQRLPRPGRPREVHIGRALGRVAPSMLLCSLSAICFP 759
Db 492 QVVPFLALGLGVDHIFMLTAAVYAESNRR--EQTKL-----ILKVGFSILFSACSTAGSPF 545
QY 760 LGALTMPAVRTFALTSGLAVILDFLLQMSAFVALLSLDSKQASRLDV-CCCV----- 813

Db	546	AAAFIPVPAKVFCLQRAIVMCSNLAALVFPAMISLDLRRRTAGRADIFCCCFPWKE	605	PA	(GEMM/) GEMMILL R M.
Qy	814	KQOELPP-----PQOGEGL-----LLGF	831	XX	(DRAB/) DRABKIN H A.
Db	606	QPKVAPPVPLNNNGRGARHPKSCNNRVPLPAQNPLLEQORADIPGSSHSLASFSLATP	665	XX	Gemmill RM, Drabkin HA;
Qy	832	FQKAYAPPELLHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSVLLDYFLPLN	891	XX	WPI; 2002-712395/77.
Db	666	AFQHYTPFLMRSWVKFLTMGFLAALISLVASTRLODGLDIIDLVPKDSNEHKFLDAQT	725	XX	Novel Translocation in Renal cancer from Chromosome 8 genes, useful for
Qy	892	RYFEVGAAPVYFTTILGYNFSSEAGNNAICSSAGCNCNFTOKIOVATEFPQSYLAIP--	949	XX	detection of tumors, comprises rearrangements in the t(3;8)(p14.2;q24.1)
Db	726	RLF--GFYSMAVTQG-----NFEYPTQQQLRDY-HDSFVRPHV	763	XX	chromosomal translocation which occurs in renal and thyroid carcinomas.
Qy	950	-----ASSWDDFDIDLWTPSSCCRLYISGPNKDKFCPSTVNSLNCNKC-----	993	XX	Example 1; Fig 2C; 49pp; English.
Db	764	IKNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWPNASSD	811	XX	The invention relates to an isolated TRC8 (Translocation in Renal cancer
Qy	994	-----MSITWGSVRPSVEQ-----FHKYL-PWFNDRPNIKCPKG	1027	XX	from Chromosome 8) nucleic acid molecule, encoding a polypeptide
Db	812	AILAYKLVQTHGVNDPVDKELVLTNRLVNSDGIINQRAFYNYLSAWATND-----	862	XX	comprising a sequence of 664 amino acids fully defined in the
Qy	1028	GLAAVSTSNLTSQGVLASRWAYHKPLKNSQD-----YTEALRAARE	1071	XX	specification and comprising a sequence located in the 5' flanking region
Db	863	-VFAYG-----ASQGKLYPEPRQYFHPQ--NEYDILKIPKSLPLVYAQMPPFYLHGLTDSQ	914	XX	to the coding region of TRC8 and a sequence which occurs in certain
Qy	1072	LAANITADLRKVPGTDPAPFV--FPYTIITNVFEQYLTILPEGLFMLSCLIVPTFAVSCL	1129	XX	sporadic renal cell carcinomas. The methods are useful for detecting the
Db	915	I-KTLIGHIRDLVSUYEGFGLPNYPSGIPFWEQYMT-LRSSLAMILACVLLAALVLVS	972	XX	presence of the TRC8 gene in a biological sample, detecting alterations
Qy	1130	LLGLDLRSLGLNLLSIVMLIVDTVGFMALWDISYNAVSLINLVSAVGSVFVSHITSF	1189	XX	to the gene, such as a 3;2 human chromosomal translocation, and fused DNA
Db	973	LLLLSWAAVVLVLSLASLAQIFGAMTLIGKLSAIPAVILLSVGMWLCP--NVLLISL	1030	XX	containing the fused site of TRC8/FHIT. A nucleic acid probe is useful
Qy	1190	ALSTKPTWLERAKETISGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLLTL	1249	XX	for detecting the 3;8 human chromosomal translocation, by contacting the
Db	1031	GPMTSVGNQRVQVLSMQSLGLPLVHGLMTSGVAVFMLSSTPFVIRHFCHLLLVLCV	1090	XX	nucleic acid probe with a biological sample to be tested, and determining
Qy	1250	GLLHGLVFLFVLISVYVGPDPNPALAEQKRAEEAAVAMVASCPNHPSRVSTADNIYNH	1309	XX	whether the nucleic acid probe specifically hybridises to the TRC8/FHIT or
Db	1091	GACNSLLVFPILLSMGPEAE-LVPLE-----HPDRISTPSPLPVRS	1131	XX	FHIT/TRC8 fusion DNA. This sequence represents a TRC8 related protein of
Qy	1310	S 1310		XX	the invention
Db	1132	S 1132		XX	Sequence 1286 AA;
RESULT 73					Query Match
ABJ10929	ABJ10929 standard; protein; 1286 AA.				6.9%; Score 474; DB 5; Length 1286;
XX					Best Local Similarity 20.6%; Pred. No. 8.5e-34;
AC					Mismatches 246; Conservative 176; Mismatches 465; Indels 306; Gaps 43;
XX					
DT	12-DEC-2002 (first entry)				
XX	TRC8 related protein Dm Ptc SEQ ID No 3.				
DE					
XX	TRC8; Translocation in Renal cancer from Chromosome 8; fused DNA; 3,2;				
KW	FHIT/TRC8 fusion DNA; sporadic renal cell carcinoma; TRC8/FHIT; TRC8FHIT;				
XX	human chromosomal translocation.				
OS	Drosophila melanogaster.				
PN	US2002106656-A1.				
XX					
PD	08-AUG-2002.				
XX					
XX	02-JUL-2001; 2001US-00898533.				
PF					
XX	12-MAR-1998; 98US-0077723P.				
PR	12-MAR-1999; 99US-00268140.				
XX					
XX					

QY 821 -----PQGBGL-----LLGPFQKAYAPF 839
 Db 614' LPLNNNGRGARHPKSCNNRVPPLPAQNPLEQADIPGSSHSLASFSLATFAFQHYTFP 673
 QY 840 LLHWITRGVVLALLFALFGVSLYSMCHISVGLDQELALPKDSVILLDYFLFLNRYPEVGAP 899
 Db 674 LMRSWKFLTVMGFLAALSSLYASTRLQDGLDIDLVPKDSNEHKFLDAQTRLF--GPF 731
 QY 900 VVFVTLTGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIP----- 949
 Db 732 SMYAVTQG-----NFEYPTQQLRDY-HDSFVRPHVKNKNGSL 771
 QY 950 ASSVDDFDIMLTPSSCCRLYISGPNKDKPCPSTVNSLCLNKC-----MS 995
 Db 772 PDFWLLLFSEWL-----GNLQIFDEEYRDGRLTKECWPPNASSDAILAYKLI 819
 QY 996 ITWGSVRPSVEQ-----FHKYL-PWFLNDRPNKCPKGLAAYSTS 1035
 Db 820 VOTGHVDNPNDELVLVTLNLSNDGIINQRAFTNLSAWATND-----VFAYG-- 867
 QY 1036 VNLTSQGVLASRFMAYHKPLKNSQD-----YTEALRAARELANITAD 1079
 Db 868 ---ASQKLYPPEQYFHP--NEYDLKTPKSLPLVYAQMPFYLHGLTDTSOI-KTLIGH 921
 QY 1080 LRKVPQTDPAFV--PPYITITNVFVEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLRS 1137
 Db 922 IRDLSVKYEGFGLPNYPGIPFWEQYMT-LRSSLAMILACVLLAALVLSLLLSVMA 980
 QY 1138 GLNLLISIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTW 1197
 Db 981 AVLIVLSLASLAQIFGANTLLGILKLSAIPAVILISVGMWLCF--NVLLISGFMTSVGN 1038
 QY 1198 LERAKEATISMGSAFVAGVAMTNLPGILVLGLAKAQIQLIFFERLMLTLTLGLLHGLVF 1257
 Db 1039 RRRVQLSQMSGLPLVHGMLTSGVAVFMLSIPPEFVIRHFCWLLLVLCVGCNSLLV 1098
 QY 1258 LPVILSYGPDVNPALALQKRAEAAVAVMVASCPNHPSPVSTADNIYVNS 1310
 Db 1099 FBILLSMVGPEAB-LVPLE-----HPRISTSPPLVRSS 1132

RESULT 74
 AAR86304
 ID AAR86304 standard; protein; 1299 AA.
 AC AAR86304;
 DT 23-MAR-1996 (first entry)
 DE Drosophila patched protein.
 KW patched protein; hedgehog receptor protein; Drosophila; sonic hedgehog;
 KW desert hedgehog; Indian hedgehog; drug screening; nervous system disease.
 OS Drosophila.
 FH Key Location/Qualifiers
 FT Domain 93..426
 FT /note= "Extracellular domain"
 FT Domain 700..966
 FT /note= "Extracellular domain"
 XX W09518856-A1.
 XX 13-JUL-1995.
 XX 30-DEC-1994; 94WO-US014992.
 XX 30-DEC-1993; 93US-00176427.
 PR 14-DEC-1994; 94US-00356060.
 XX (HARD) HARVARD COLLEGE.

(IMCR) IMPERIAL CANCER RES TECHNOLOGY.

Ingham PW, McMahon AP, Tabin CJ;

WPI: 1995-255060/33.

N-PSDB; AAQ91860.

Hedgehog-like protein(s) and nucleic acid(s) encoding them - useful to treat degenerative nervous system disorder(s) and in gene therapy.

Disclosure; Page 170-76; 210pp; English.

The sequence represents a patched protein from *Drosophila*, which acts as a receptor for *Drosophila* hedgehog protein (AAR77337). The receptor or extracellular domain subfragments may be used in a drug screening assay, along with a vertebrate hedgehog protein, e.g. sonic hedgehog, desert hedgehog and Indian hedgehog proteins (AAR77338-R77341 and AAR77343-R77345). Compounds which inhibit or potentiate complex formation between the receptor and hedgehog protein may be used in therapy of neuromuscular, autonomic or central nervous system disorders

Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 2; Length 1299;

Best Local Similarity 20.4%; Pred. No. 6.7e-33;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLGQFQCGWGTWASWPLTTLVLSVIPVVALAAGLVFELTTPV 380

Db 43 KKGAGSRITAIYLRVSFQSHLETGLSSVQKHAGKLVFVAILVSTFCVGLKSAQIHSKVH 102

QY 381 ELWASPNQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSYRYSLLGLGPKNFSGIL 436

Db 103 QLMIQEGGSLRAELAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153

QY 437 LDDLLELLELQERLHRHQVSPQAQNISLQDICIAPLNPNSTLY-----DCC 486

Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPSTPSEGIYIIEQLRHLIPCS 204

QY 487 INSLAQYQN-----NRTLALLTAN-----QTLMGQTSQVMDKDH 521

Db 205 IITPLDPCWEGSQLGPESAVVPIGLNORLLMTLNPASVMQMKQKSEKISPDFETV 264

QY 522 FLYCANAPLTFKDGTAALSCM-----ADYGAPVPPFLAIG--GYK 560

Db 265 EQYMGRAIA---SGYMEKPCPLNPNCPDPAPKNSTQPPDVG-----ILSGGCGYA 317

QY 561 GKDYSEABALIMTFLSNYPAGDPLQA-----589

Db 318 AKHMHPBELIVGGAARN-RSGHLRKAQALQSVQVQLMTEKEMYDQWQDNKYVHHLGWTQE 376

QY 590 -----KLWEERAF---LEEMRAFORMAGMPQVTTAERSLEDEINRTTAEDLPFATSY 640

Db 377 KAAEVLNAWQNRFSREVSQLRKQRIATNYDIYVFSSAALDDILAKFSHPALSIVIGV 436

QY 641 IVIFLYISLALGSYSWSRVVDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQ 700

Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGAGVLLMCFSTAAGLSALLGIVFNAATQ 492

QY 701 VVPFLVLSVGADNIFVLEVQRLPRRPGEPREVIHGRALGRVAFSMLLCSISEAICPFL 760

Db 493 VVPFLALGLGVDDHIFMLTAAYAESNRR--EQTKL-----ILKKVGP SILFSACSTAGSPFA 546

QY 761 GALTTPMPAVRTFALTSLGLAVILDFFLOMSAFVALLSLDSKQREASRLDV-CCCV-----K 814

Db 547 AAFIEVPALKVFCLOAAIVMCSNTAAALLVFPAMTSLDLRRRTAGRADIFCCCFPMWKEQ 606

QY 815 POELPP-----PQGBGL-----LLGPF 832

Db 607 PKVAPVLPVLPNNNGRGARHPKSCNNRVALPAQNPLEQADIPGSSHSLASFSLATFA 666

QY 833 QKAYAPFLHMITRGVILLFLALFGVSLYSMCHISVGLDQELALPKDSVILLDYFLFLNMR 892

Db 667 FOHTPTFLMRWVFLTWGFLAALISSYASTRLQDGLDIIIDLPVKDSNEHKFLDAQTR 726
Qy 893 YFEVGAPYFVTTLGYNFSSBAGNAICSSAGCANNFSTQKIQYATFPEQSYLAIP--- 949
Db 727 LF--GFYSWYAVTQG-----NFEYPTQQQLRDY-HDSFVRVPHVI 764
Qy 950-----ASSWVDDFIDMLTPSSCCRLYISGPNKDKFCPSVNSLNCIKNC----- 993
Db 765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGLTKECWFFNASSDA 812
Qy 994-----MSITMGSVRSVRSQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db 813 ILAYKLIYQTHVDNPNVDKELVTLNRLVNSDGIINORAFYNLSAW-----ATNASSPEL 868
Qy 1029 LAA--YSTSVNLTSQGVLASRFMAYHKPLKNSQD-----YTRALRAAR 1070
Db 869 LRANCIRNRANGASQGLPYEPQVQHQP--NEYDLKIPKSLPLVYAQMPPYLHGLTDT 926
Qy 1071 ELAANITADLRKVPGTDAFEV--FPYITVNVFYEOYLTILPEGLFMLSCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDSVKYEGFGLPNYPSPGIPFIFWEQYMT-LRSSLAMTILACVLLAALV 984
Qy 1129 LILGLDRLSGLNLALSIVMILVDVTGFMALWDISYNAVSLINLVSAGVSVFVSHITRS 1188
Db 985 SULLSVWNAVILSVLASIAQIFGAMTLLGKLSAPVAILLSVGMCLCF--NVLS 1042
Qy 1189 FAISTKPTWLREAKATISMSGSAVAGVAMTLPGLVILGLAKAQIQLFPRMLNLT 1248
Db 1043 LGFMTSVGNRRQRRVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFVFIRHFCWLLV 1102
Qy 1249 LGLLGLVLPVILSVGPDVNPALALEQKRAEEAAVAMVASCNPHSPRSTADNIYN 1308
Db 1103 VGACNSLLVFPILLSVMGPEAE-LVPLE-----HPRISTSPSLPVR 1143
Qy 1309 HS 1310
Db 1144 SS 1145

RESULT 75
ID AAB05389 standard; protein; 1299 AA.
XX AAB05389;
XX 12-SEP-2001 (first entry)
XX Human patched protein.
XX Human; hedgehog; morphogenic signal; neuron; patched protein;
KW embryonic patterning; cell culture; cell differentiation; ischaemia;
KW cell proliferative disorder; intracerebral grafting; Huntington's chorea;
KW neurological disorder; Alzheimer's disease; Parkinson's disease;
KW anyotrophic lateral sclerosis; ALS; multiple sclerosis.
XX Homo sapiens.
OS US6261786-B1.
XX 17-JUL-2001.
XX 02-JUL-1996; 96US-00674509.
XX 30-DEC-1993; 93US-00176427.
XX 14-DEC-1994; 94US-00356060.
XX 04-MAY-1995; 95US-00435093.
XX 05-JUN-1995; 95US-00460900.
XX 05-JUN-1995; 95US-00462386.
XX (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
PA (HARD) HARVARD COLLEGE.
XX Marigo V, Tabin CJ, Ingham PW, McMahon AP;

XX WPI; 2001-440859/47.
DR N-PSDB; AAD10170.
XX
PT Screening compounds that potentiate or inhibit binding of hedgehog
PT polypeptide to naturally occurring patched receptor, comprises contacting
PT polypeptide with receptor and test compound, and detecting change in
PT binding.
XX
PS Disclosure; Col 193-202; 127pp; English.
XX
XX The present invention relates to assay for screening compounds that
XX potentiate or inhibit binding of hedgehog polypeptide to naturally
CC occurring patched receptor. The hedgehog proteins comprise morphogenic
CC signals produced by embryonic patterning centres, and are involved in the
CC formation and maintenance of ordered spatial arrangements of
CC differentiated tissues in vertebrates, both adult and embryonic. The
CC vertebrate tissues both in vitro and in vivo. The invention also relates
CC to a method for modulating growth, differentiation or survival of a
CC mammalian cell (e.g. neuron, testicular cell) responsive to hedgehog
CC induction. Hedgehog agonists and antagonists can be used in cell culture
CC techniques to enhance survival and maintenance of neurons and various
CC vertebrate organogenic pathways. The hedgehog gene is useful in
CC determining whether a patient is at the risk of disorder characterised by
CC unwanted cell proliferation or aberrant control of differentiation. The
CC hedgehog proteins or mimetics can be used to induce foetal neurons
CC especially neuronal stem cells in intracerebral grafting. The protein or
CC its mimetic can be used in the treatment of neurological conditions e.g.
CC injury to nervous system, ischaemia resulting from stroke, Alzheimer's
CC disease, Parkinson's disease, Huntington's chorea, anyotrophic lateral
CC sclerosis (ALS) and multiple sclerosis. The present sequence is human
CC patched protein related to the invention. Human patched protein is
CC provided in screening assays as a whole protein, or alternatively as a
CC fragment of the full length protein which binds to hedgehog proteins. The
CC patched protein can also be provided as a glycosylated protein
XX
SQ Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 4; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
Qy 321 KGTSLSDKLSFSTHTLLGQFFQGMGTWVSWPLTILVLSPVVALAGLVTELTTPV 380
Db 43 KGKARGRTAIYLRVSFQSHLETLLGSSVQKAGKVLFAILLVSTFCVGLSAQIHSKVH 102
Qy 381 ELWSAPNSQARSEKAFHQHFG-PFRTNQVILTA--PNRSSRYVDSLLGLPKNFGSL 436
Db 103 QLWIQEGGGLEAELAYTKTIGEDSATHQLLIQTHDFNASVLHPQALL-----A 153
Qy 437 DLDLLELLELQERLRHLQVWSPEAQRNLSQDICVAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMSTPSPFEGIIYIEQILRHLP 204
Qy 487 INSLLIQYFQN-----NRTLLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVPGINQRLLTINLPASVMQMKQKMESEKISDFETV 264
Qy 522 FLYCANAPLTFKDGITALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EYMKRAAJA---SGYMEKPCLNPNPCPDAPNKNSTQPPDVGA-----ILSGGCGYA 317
Qy 561 GKDYSEAEALIMTFSLNNYPAGDPRLQAQ-----589
Db 318 AKHWHPEELIVGGAKEN-RSGHLRKAQALQSVVQLWTEKMYDQNDNVKVVHLLGWTQE 376
Qy 590-----KLWEEAF---LEENRAFQRMAGMFQVTFATERSLEDEINRTTAEDLPFATS 640
Db 377 KAAEVLNMQRNFRSREVEQLRKRQRIATNYDIYFSSAALDDILAKFHSPLSALSVIGV 436
Qy 641 IVTFILYISLALGYSYSSSRVWVDSKATLGLGGVAVVLGAVMAAMGFFSLGIRSSLVILQ 700

Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAAGLGLSALLGLVFNASTQ 492
 QY 701 VVPELVLSVADNIFIVLEVYQRLPRPGEPREVIHGRALGRVAPSMMLCSLSEAICFFL 760
 Db 493 VVPELVLSVADNIFIVLEVYQRLPRPGEPREVIHGRALGRVAPSMMLCSLSEAICFFL 760
 QY 761 GALTMPAVRTALTSLGLAVILDIFLLOMSAFVALLSLDSKREASRLDV-CCCV-----K 814
 Db 547 AAFIPVPALKVFCLOAAIVMCSNLAALVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
 QY 815 POELPP-----PQGBGL-----LLGPF 832
 Db 607 PKVAPVPLPNNNGRGARHPKSCNNRVLPQNPPLLEQRADIPQSSHSLSASFSLATPA 666
 QY 833 OKAYAPFLHWTGRGVLLFLALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLFLNR 892
 Db 667 FQHYTPFLMRSWKFLVTMGFLAALLISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
 QY 893 YFEVGAPVYVTVLGNFSEBAGMAICSSAGCNPFSTQKIQYATEPPEQSYLAIP--- 949
 Db 727 LF--GFYSMYAVTQG-----NFEYPTQQQLRDY-HDSFVRVPHVI 764
 QY 950 -----ASSWDDRIDMTLTPSSCCRLXISGPNKDKFCFSTVNSLNCNKC----- 993
 Db 765 KNDNGLPDFWLLFSEWL-----GNLQKIFDEEYRDGRLTKCEWFPNASSDA 812
 QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRNPKCPKGG 1028
 Db 813 ILAYKLIQVQTHVNDPVDKELVITNRLVNSDGIINQRAFYNYLSAW-----ATNASSPTBL 868
 QY 1029 LAA--YSTSVNLTSGQVLAASFMAVHKPLKNSQD-----YTEALRAAR 1070
 Db 869 LRANCIRNRANGASQCKLYPEPRQVPHQF--NEYDLKIPKSLPLVYAQMPFYLHGLTDT 926
 QY 1071 ELAANITADLRKVGCTDPAFEV--PPTITNVFYQYLTILPEGLMSLCLVPTFANVC 1128
 Db 927 QI-KTLIGHIRDSLVKYEGLNTPSPGIPFIFWESQYMT-LRSSLAMILACVLLAALVLV 984
 QY 1129 LLLGLDLSGLLNLASIVMLVDTVGFMAIWDISYNAVSLINLNAVSGMSVFEVSHITRS 1188
 Db 985 SLLLSVWAVALVILSVLASLAQIFGAMTLLGLKLSAIPAVILILSVGNMLCF--NVLS 1042
 QY 1189 PAISTKPTWLERAKATISMGSAGFAGVAMTNLPGILVILGLAKAQLIQIFPFLMLLITL 1248
 Db 1043 LGFMTSVGNRRQRRVQLSMQMSLGLPLVHGMLTSGVAVFMLSTSPFBEFVIRHFCWLLLVLC 1102
 QY 1249 LGLLGLVPLVILSVGPDVNPALALEQKRAEEAAVAVMWASCPNHPRSVSTADNIYN 1308
 Db 1103 VGACNSLLVFPILLSWVGPEAB-LVPLE-----HPDRISTPSPLPVR 1143
 QY 1309 HS 1310
 Db 1144 SS 1145
 RESULT 76
 ID AAB85751
 XX AAB85751 standard; protein; 1299 AA.
 AC AAB85751;
 DT 29-OCT-2001 (first entry)
 XX Hedgehog (hh) polypeptide related sequence Id No. 48.
 XX Hedgehog protein; sonic hedgehog; Shh; indian hedgehog; Ihh; Dhh;
 KW desert hedgehog; cell differentiation.
 XX Unidentified.
 XX US6271363-B1.
 XX 07-AUG-2001.

XX 20-OCT-1997; 97US-00954698.
 XX 30-DEC-1993; 93US-00176427.
 PR 14-DEC-1994; 94US-00356060.
 PR 04-MAY-1995; 95US-00435093.
 PR 05-JUN-1995; 95US-00462386.
 XX (HARD) HARVARD COLLEGE.
 PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
 XX Ingham PW, McMahon AP, Tabin CJ;
 XX WPI; 2001-456723/49.
 DR N-PSDB; AAH76131.
 XX Novel nucleic acid encoding a hedgehog polypeptide, used to produce the
 PT polypeptide, which is used to promote proliferation, survival, and/or
 PT differentiation of neuronal and mesodermal tissue.
 XX Disclosure; Col 181-188; 118pp; English.
 XX The invention relates to nucleic acids encoding hedgehog proteins
 CC selected from sonic hedgehog (Shh), Indian hedgehog (Ihh), Desert
 CC hedgehog (Dhh) polypeptides. The hedgehog genes are involved in the
 CC formation of ordered spatial arrangements of differentiated tissue in
 CC vertebrates. The nucleic acid sequences are useful for producing hedgehog
 CC proteins, used for promoting differentiation of, or survival of
 CC differentiated, neuronal cells, and for promoting proliferation, survival
 CC or differentiation of mesenchymal, endodermal or ectodermal tissue,
 CC particularly chondrocytes, or testicular germ line cells. The present
 CC sequence represents a hedgehog polypeptide related sequence
 XX Sequence 1299 AA;
 SQ
 QY Query Match 6.7%; Score 464.5; DB 4; Length 1299;
 Db Best Local Similarity 20.4%; Pred. No. 6.7e-33;
 QY Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
 Db 321 KGTSLSDKLSPTHTLLGQFGGCTWVASWPLTILVLSVIPVVALAAGLVFTLTDV 380
 QY 43 KKGAGSRTAIYLRVSFQSHLETGLSSVQKHAGKLVFVAILVLSFTCVGLKSAQIHSKVH 102
 Db 381 ELWASNPQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSRYDSLLGLPKNPSGIL 436
 QY 103 QLMIQEGGGLAEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
 Db 437 DLDLLELLELQERLHLOVMSPEAQRNISLODICVAPLNPDNTSLY-----DCC 486
 QY 154 HLEVIVKATAVKVHLVDTE-W-----GLRDMCNMPSPTSPFEGIYIYEQILRHLIPCS 204
 Db 487 INSLLOYFQN-----NRTLILLTAN-----QTLMGQTSQVMDKDH 521
 QY 205 IITPLDPCFWEGSQLGPESAVVPIGLNQLRLTTLNLPASVMQYMKQMSSEKISDFETV 264
 Db 522 FLYCANAPLTPKDGTAALALSCM-----ADYGAPVPPFLAIG--GYK 560
 QY 265 EGYMKRAAIA---SGYMEKPCLNPLNPNCPDAPNKNSTQPPDVG-----ILSGCYGVA 317
 Db 561 GROYSEAEALIMTFSLNYPAGDPLAQA-----589
 QY 318 AKHMWPEELIVGGAQRN-RSGHLRKAQALQSVVQLMTEKEMYDQMDQNKVHHLGWTQE 376
 QY 590 -----KLWEAP---LEEMAFORMMAGMPQVTTAERSLEDEINRTTAEDLPFATSY 640
 Db 377 KAAEVLNAWQRNFSREVSQLRKQRIATNYDIYVFSSAALDDILAKFSHPALSIVIGV 436
 QY 641 IVIFLYISIALGSSYSSWSRVVDSKATLGLGVAVVTLGAVMAAGFFSVYIGRSSIVLIQ 700
 Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAAGLGLSALLGLVFNASTQ 492
 QY 701 VVPELVLSVADNIFIVLEVYQRLPRPGEPREVIHGRALGRVAPSMMLCSLSEAICFFL 760

Db 493 VVPFLALGLVDHIFMLTAAVAENRR--EQTKL-----ILKKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTFALTSGVLDFELLOMSAFVALLSLDSKQEAASRLDV-CCCV-----K 814
Db 547 AAFIPVPALKVFCLOAAIVMCSNLAALLVFPAMISLDLRRTAGRADIFCCCFPVWKEQ 606
QY 815 POELPP-----PGQEGIL-----LIGFF 832
Db 607 PKVAPPVLPNNNGRGARHKPCNNNRVALPAQNPLEQRAADIPGSSHSLASPSLATFA 666
QY 833 QKAYAPFLHWTTRGVLLFLALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLPLNR 892
Db 667 FQHYTFPLMRWVKFLTWGFLAALISSIYASTRLQDGLDIIDIVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYFVTLTYNFSSEAGNMAICSSAGCANNFSFTQIKIYATEPEQSYLAIP--- 949
Db 727 LF--GFYSNYATQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDFIDWLTSSCCRLYISGNKDKFCFSTVNSLNCIKNC----- 993
Db 765 KNDNGGLPDFWLLLESEWL-----GNLQKIFDEEYRDGRLTKECWFNPASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db 813 ILAYKLIVQTGHVDPVDPVKELVLTWRLVNSDGIINQRAFYNYLSAW-----ATNASSPTEL 868
QY 1029 LAA--YSTSVNLTSDQVLAASFMAHYKPLKNSOD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLYPEPRQYFHP--NEYDLKIPKSLPLVYAQMPPYVHLGLTDTS 926
QY 1071 ELAANITADLRKVPGTDPAFEV--FPYITTNVYFQYLTILPEGLFMLSCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDLSVKYEGFGLNFPSPGIPFIFWEQYMT-LRSSLAMILACVLLAALVIV 984
QY 1129 LLLGLDRLSGLANLLSIVMLVDTVGFMAWDISYNAVSLMLVSAVGSVFEVSHITRS 1188
Db 985 SLLLSVMAAVILVSVLASLAQIFGAMTLLGIKLSAPAVILLISVGMMLCF--NVLIS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVFAGVAMTLPGLVILGLAKAQIQTFFPLMLLITL 1248
Db 1043 LGFMTSVGNRRQRRVQSLMQMSLGPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
QY 1249 LGLLHGLVFLPVLTSVYGVDPVNPALALEQKRAEEAAVAVMWASCPNHPRSVSTADNIYN 1308
Db 1103 VGACNSLLVFPILLMSVGPAAE-LVPLE-----HPDIRSTPSPLPVR 1143
QY 1309 HS 1310
Db 1144 SS 1145
RESULT 77
AAB31232
ID AAB31232 standard; protein; 1299 AA.
XX
AC AAB31232;
XX
DT 20-APR-2001 (first entry)
XX
DE A consensus hedgehog-related protein.
XX
KW Hedgehog related-protein; sonic hedgehog protein; Shh; ischemia; stroke;
KW desert hedgehog protein; Dhh; indian hedgehog protein; Ihh; neuron;
KW neurological condition; nervous system injury; tumour-induced injury;
KW aging; Alzheimer's disease; chronic neurodegenerative disease;
KW Parkinson's disease; Huntington's chorea; amyotrophic lateral sclerosis;
KW spinocerebellar degeneration; chronic immunological disease;
KW multiple sclerosis.
XX
OS Synthetic.
OS Unidentified.
XX
PN US6165747-A.

XX PD 26-DEC-2000.
XX PF 05-JUN-1995; 95US-00460900.
XX PR 30-DEC-1993; 93US-00176427.
XX PR 14-DEC-1994; 94US-00356060.
XX PR 04-MAY-1995; 95US-00435093.
XX PA (HARD) HARVARD COLLEGE.
XX PA (IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX PI Ingham PW, McMahon AP, Tabin CJ, Marti-Gorostiza E, Bumcrot DA;
XX DR WPI; 2001-079847/09.
XX DR N-PSDB; AAC87096.
XX PT Polynucleotides encoding hedgehog proteins, useful for treating diseases
XX PT of nervous system such as Alzheimer's disease, Parkinson's disease,
XX PT Huntington's chorea, amyotrophic lateral sclerosis, multiple sclerosis.
XX PS Disclosure; Col 183-192; 119pp; English.
XX CC The present sequence represents consensus a hedgehog related-protein. The
XX CC specification describes a sonic hedgehog protein (Shh), a desert hedgehog
XX CC protein (Dhh), and an indian hedgehog protein (Ihh). The hedgehog
XX CC polynucleotides are useful in diagnostic, in antisense therapy and in
XX CC therapeutic assays for detecting and treating disorders involving, e.g.,
XX CC aberrant expression of vertebrate hedgehog homologue. Hedgehog
XX CC polypeptides are useful therapeutically to enhance survival of neurons
XX CC and other neuron cells and in treating neurological conditions deriving
XX CC from acute, subacute, or chronic injury to the nervous system, including
XX CC traumatic injury, chemical injury, vascular injury and deficits (such as the
XX CC ischemia resulting from stroke), together with infectious/inflammatory
XX CC and induced-induced injury, aging of the nervous system including
XX CC Alzheimer's disease, chronic neurodegenerative diseases of the nervous
XX CC system, including Parkinson's disease, Huntington's chorea, amyotrophic
XX CC lateral sclerosis, spinocerebellar degenerations, and chronic
XX CC immunological diseases of the nervous system or affecting the nervous
XX CC system, including multiple sclerosis
XX SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 4; Length 1299;
Best Local Similarity 20.4%; Pred. NO. 6.7e-33;
Matches 245; Conservative 174; Mismatches 479; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSFTHTLLQPFQCGWGTWASWPLTILVLSVIPVVALAAGLVFTELTDPV 380
Db 43 KKGKRGSRTAIYLRSVFQSHLETGSSVQKHAGKVLFAILLVSTFCVLGKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFG-PFFRTNQVILTA---PNRSYRYSDLSLLGPKNPSGIL 436
Db 103 QLWIOEGGGLAEALAYTOKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLRLHQVMSPEAQRNISLQDICYAPLNPDNTSLY-----DCC 486
Db 154 HLEVLKATAVKVHLYDTE-W-----GLRDMCNMPSPTSPGFIYIEQILRLHIPCS 204
QY 487 INSLIQQYFQN-----NRTILLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVIPGLNQRLLTTLNPASVMQMKQSEEEKISPDFETV 264
QY 522 FLYCANAPITFDGKTALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EQMKGAALIA---SGMYKPCPLNPLNPNCPDTPAPKNKSTQPPDVGA---ILSGGCGYVA 317
QY 561 GKDYSEAEALIMTFSLNPNYPAGDPRLAQA----- 589
Db 318 AKHMHWPPEELIVGGAQRN-RSGHLRKAQALQSVVOLMTEKEMYDQMDNQVKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFORMAGMCFQVTFATERSLEDEINRTTAEDLPIFATSY 640

Db 377 KAAEVLNANQRFSEVEQLLRKQSRATNYDIYVSSAALDDILAKFHPGALSIVIGV 436
 QY 641 IVIFLYISALGSYSWSRVMDSKATLGLGGVAVVVLGAVMAAMGPFYSYLGRSSIVLIQ 700
 Db 437 AVTVLY--AFCTLLRW-RDPRGQSSVGAVGLLMCFSTAGLGLSALLGIVFNAASQ 492
 QY 701 VVPFLVLSGADNIFIVLEYQRLPRPQBPBRVHIGRALGRVAPSMMLCSISEAICPFL 760
 Db 493 VVPFLALGLGVDFHFMFTAAYAESNR--EQTKL-----ILKVGPSILFSACSTAGSFFA 546
 QY 761 GALTMPAVRTALTSLGVLIDELQMSAFVALLSLDSKROEASLDV-CCCV-----K 814
 Db 547 AAFIPVPAKVFCLOAAIWMCSNLAALVFPAMISLDRRRRAGRADIFCCCFPVWKEQ 606
 QY 815 PQELPP-----PQGEGL-----LLGFF 832
 Db 607 PKVAPPVLNNNGRGARHPKSCNNRVALPAQNPLEQRADIPGSSHSLASFSLATPA 666
 QY 833 QKAYAPFLHWTTRGVVLLFLALFCVLSYMSCHISVGLDQELALPKDSYLLDYFLNLR 892
 Db 667 FQHYTPFLMRSWKFLTVMGFLAALISSLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
 QY 893 YFEVGAPVFTVTLGYNFSSSEAGMNAICSSAGCNPSEFTQKIQYATEPPEQSVLAIP--- 949
 Db 727 LF--GFYSIAVTQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
 QY 950 -----ASSWDDFDIMLTPSSCCRLYISGPNKDKFCPSVTNSLNLKNC----- 993
 Db 765 KNDNGGLPDFWLLFSEWL-----GNLQKIPDEYRDGRLTKECFPPNASSDA 812
 QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRNRIKCPKGG 1028
 Db 813 ILAYKLIVQTHVDNPDVKELVLNRLVNSDGLINQAFNYLSAW-----ATNASSPTL 868
 QY 1029 LAA--YSTSVNLTSQVLAISRFMAVHKPKNSQD-----YTEALRAAR 1070
 Db 869 LRANCIRNRANGASQKLYPEPRQYHPQ--NEYDLKIPKSLPLVYAQMFPYLHLGLTDS 926
 QY 1071 ELAANITADLRKVPGTDPAPFV--FPYITNVFYEQYLTPLEGFLMLSCLVPTFAVSC 1128
 Db 927 QI-KTLIGHIRDLVKYEGFGLPNYPSGIPFIFWEQYMT-LRSSLAMILACVLLAALVLV 984
 QY 1129 LLGLDLRGLNLLSIVMLVDTVGFMAWMDISYNAVSLINLVASGVMSVFSVSHITS 1188
 Db 985 SLLLSVMAAVLVLSLASLAQIFGAMTLGLKLSAIPAVILISVGNMCLCF--NVLTS 1042
 QY 1189 PAISIKPTWLERAKEATISGSAVPAVAMTNLPGTLVLGLAKAQIQTFFRLNLLTL 1248
 Db 1043 LGFMTSVGNRRQVQLSMQMSLGPLVHGLMTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
 QY 1249 LGLLHGLVLPVILSVYGVDPNPALAEQKRAEEAAVAVMWASCPNHPRSVSTADNIYN 1308
 Db 1103 VGACNSLLVPFILLSVNGFEAR-LVPLE-----HPRISTPSPLPVR 1143
 QY 1309 HS 1310
 Db 1144 SS 1145

RESULT 78

ABB79150

ID ABB79150 standard; protein; 1299 AA.

XX ABB79150;

XX ABB79150;

XX 06-AUG-2002 (first entry)

XX Patched protein sequence SEQ ID NO:42.

XX

XX

KW Sonic hedgehog; Shh; desert hedgehog; Dhh; Indian hedgehog; Ihh;

KW antiparkinsonian; antiarrhythmic; neuroprotective; anticonvulsant;

KW cytostatic; neurotropic; spermatogenesis; peripheral nervous system;

KW central nervous system; Alzheimer's disease; Parkinson's disease;

KW

KW Huntington's disease; arrhythmia; nerve degeneration; multiple sclerosis;
 immunological disorder; neoplastic; hyperplastic.

Unidentified.

US6384192-B1.

07-MAY-2002.

20-OCT-1997; 97US-00957874.

30-DEC-1993; 93US-00176427.

14-DEC-1994; 94US-00356060.

04-MAY-1995; 95US-00435093.

05-JUN-1995; 95US-00462386.

(HARD) HARVARD COLLEGE.

(IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.

Ingham PW, McMahon AP, Tabin CJ;

WPI; 2002-442817/47.

N-PSDB; ABN87568.

New vertebrate hedgehog-related proteins, useful e.g. for promoting

differentiation, survival and proliferation of cells, e.g. for treating

neurodegeneration.

Disclosure; Col 171-180; 116pp; English.

The present invention describes an isolated and/or recombinant

polypeptide (I) comprising a hedgehog (hh) amino acid (aa) sequence

encoded by a nucleic acid (II) that hybridizes under stringent conditions

to 1 of 6 sequences (see ABN87544, and ABN87546 to ABN87550). (I) binds

to a natural patched receptor. Specifically claimed example of (I) are

given in ABN79132 and ABN79134 to ABN79138. (I) has antiparkinsonian,

neurotropic, neuroprotective, anticonvulsant, antiarrhythmic and cytostatic

activities. (I) induces the expression of the BMP-2 and -4 genes, and of

the Hoxd gene. (I) can be used: (i) to promote differentiation of

neuronal cells and survival of the differentiated cells, specifically

dopaminergic or motor neurons, proliferation of chondrocytes, and

proliferation, differentiation and/or survival of mesodermal or

ectodermal cells, either in cell cultures (particularly for preparation

of transplants) or therapeutically; (ii) for detecting loss of response,

in tissues or, to hh proteins; (iii) in drug screening (to identify

(ant)agonists, useful e.g. for inhibition of spermatogenesis); and (iv)

for isolation of cognate receptors. (I) may be used therapeutically to

treat e.g. injuries/defects in the central or peripheral nervous systems,

including Alzheimer's, Parkinson's and Huntington's diseases, or

arrhythmias caused by nerve degeneration; immunological disorders of the

nervous system, e.g. multiple sclerosis; neoplastic and hyperplastic

alterations in the central nervous system, also to promote attachment of

protheses. The present sequence represents a patched protein, which is

used in the exemplification of the present invention

Sequence 1299 AA;

Query Match 6.7%; Score 464.5; DB 5; Length 1299;

Best Local Similarity 20.4%; Pred. No. 6.7e-33;

Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

QY 321 KGTSLSDKLSFSTHTLLGQFFQCGMTWASWPLTILVLSVFWALLAAGLVTELTDPV 380

Db 43 KKGARGRTAIVLRVFSQSHLETSSVQKXGKGVLFVAILVLSFCVGLKSAQHSKVH 102

QY 381 ELWASPNQARSEKAFPHQHFQ-PFFRTNQVILTA---PNRSSRYDSLLGPKNFSGIL 436

Db 103 QLWIQEGGGLAEALVYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153

QY 437 DLDLLELLELLEQLERLHLQVMSPEAQRNLSLODICVAPLNPNTSLY-----DC 486

Db 154 HLEVLVKAIVAKVHLVDYDE-W-----GLRDMCNMPSFPFEGIIYIEQILRLHLP 204


```
CC invention.
XX
SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 6; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSKLSFSTHTLLGPOFGQGTWASWPLTILVSVVVALAAGLVFTLETATDPV 380
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 43 KGKARGSRATYILRSVFQSHLETLSGSSVQKHAGLVFVAILVLSTFCVGLKSAQIHSKVH 102
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 381 ELWASPNSQARSKAFHQHFG-PFRTNQVILTA---PNRSSRYVDSLLGPNFSGIL 436
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 103 QLMIOEGGGLAEALAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 437 DLDLLELLELOERLRHLQVMSPEAQNRISLODICVAPLNPDNTSLY-----DCC 486
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 154 HLEVLVATAVKVLHYDTE-W-----GLRDMCMNPSTSPFEGYIYIEQILRHLPICS 204
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 487 INSLLOYFQN-----NRTLLLLITAN-----QTLMGQTSQVDWDKH 521
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 205 IITPLOCWEGSQLLGPESAVVITPGLNQRLTTLNPASVMQYMKQKMSSEKISPDFEIV 264
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 522 FLYCANAPLTFKDGKTALALSCM-----ADYGAPVPFPFLAIG--GYK 560
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 265 EQMKRAAIA---SGYMEKPCPLNPLNCPDTPAPKNKSTQPPDVGA-----ILSGGCGYA 317
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 561 GKDYSEAEALIMTFSLNNYPAGDPRLAQA-----589
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 318 AKHMHWPESLIVGAKRN-RSGHLRKAQALQSVVQLMTEKEMYDQMDNYKVHHLGWTOE 376
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 590 -----KLWEEAF---LEEMRAFQRRMAGMFQVTFPAAERSLEDEINRRTAEDLPIPATSY 640
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 377 KAAEVLNAMQRNFSREVEQLLRQSRATNYDIYVFPSSAALDDILAKFSPSALSIVIGV 436
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 641 IVIFLYISALGYSNSWRVMSKATLGLGGVAVVLGAVMAAMGPFYSIGRISLVILQ 700
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGAVLLCMCFSTAAGLGLSALLGIVFNAASTQ 492
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 VVPFLVLSVADNIFIVLEYQRLPRPGEPREVHTGRALGRVAPSMILCSLSEACFPL 760
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 493 VVFLAIGLVGDHIFMLTAAYAESNR--EQTKL-----ILKRVGPSILFSACSTAGSPFA 546
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 761 GALTMPAVRTTALTSLGLVLDLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 547 AAFIPVPALKVFLQAAIVMCSNLAALVFPFAMISLDLRRRTAGRADIFCCCFPWKQ 606
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 815 POELPP-----PQCGEGL-----LLGPF 832
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 607 PKVAPPVPLNNNGRGARHPKSCNNNRVALPAQNPLLEQRADIPGSSHSLASPSLATPA 666
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 833 QKAYAPFLHWTTRGVVLLFLAFVLSYSCHISVGLDQELALPKDSYLDYFLFLNR 892
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 667 FQHYTFPLMRWVKFLTVMGFLAALISLSYASTRLQDGLDIDLVPKDSNEHKFLDQAQR 726
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 893 YFEVGAUVVFTVTLGYNFSSSEAGMNAICSSAGCNPFSTQKIYATPEPQSVLAIP--- 949
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 727 LP--GFYSYVATQG-----NFEYPTQQQLNDY-HDSFVRVPHVI 764
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 950 -----ASSWVDDFDIMLTPSSCCRLYISGPNKKPCPSTVNSLNLKNC----- 993
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 765 KNDNGGLPFWLLLFSEWL-----GNLQKIPDEYRDGRLTKECWFPNASSDA 812
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 994 -----MSITWGSVRPSEVQ-----FHKYL-PWFLNDRNRNICKPKGG 1028
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 813 ILAYKLIVQTHGVNDPVDKELVLTNRLVNSDGIINQAFYNLSAW-----ATNASSPTL 868
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1029 LAA--YSTSVNLTSQVLSARFMAVHKPKNSQD-----YTEALRAAR 1070
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 869 LRANCINRANASQSKLYPEEPQVPHQP--NEYDLKPKSLPLVVAQMPFVHLGJLTDTS 926
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1071 ELAANITADLRKVPDTPAFEV--FPYTTITNVFEQYLITLPEGLFMLSCLVPTFAVNSC 1128
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 927 QI-KTLIGHIRDLISVKYEGFLPNYPSPGIPFIWEQYMT-LRSSLAMILACVLLAALVLV 984
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1129 LLLGLDLRSGLANLLSIVMLLVDTVCGPMALWDISYNVSLINLAVSAGVMSVEFVSHITRS 1188
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 985 SLLLSVWAARVILSVLASLAQIFGAMTLLGKLSAIPAVILILSVGMMMLCP--NVLLIS 1042
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1189 FAISTKPTWLERAKETATISMSAVFAGVAMTNLPGLVLGLAKAQIQTIPFFRLNLLITL 1248
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1043 LGFMTSVGNRRQVRVQLSMQMSLGPLVHGLMTSGVAVPMLSTSPPEFVIRHFCWLLLVLC 1102
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1249 LGLLHGLVFLPVILSVGPDVNPALAEOKRABEAAVAVMWASCPNHPRSVSTADNIYNV 1308
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1103 VGACNSLLVFPILLSMVGPEAE-LVPLE-----HPDRISTPSPLPVR 1143
|| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1309 HS 1310
DB 1144 SS 1145
RESULT 80
ADD25309
ID ADD25309 standard; protein; 1299 AA.
XX AC ADD25309;
XX XX 15-JAN-2004 (first entry)
XX DE Hedgehog associated protein #3.
XX KW hedgehog; patched receptor; spermatogenesis inhibition;
XX KW ovary function inhibition; embryogenesis;
XX KW differential tissue maintenance.
XX OS Unidentified.
XX XX US6576237-B1.
XX XX 10-JUN-2003.
XX XX 16-AUG-2000; 2000US-00639695.
XX XX 30-DEC-1993; 93US-00176427.
XX PR 14-DEC-1994; 94US-00356060.
XX PR 04-MAY-1995; 95US-00435093.
XX PR 05-JUN-1995; 95US-00460900.
XX PA (HARD ) HARVARD COLLEGE.
XX PA (IMCR ) IMPERIAL CANCER RES TECHNOLOGY LTD.
XX XX Ingham PW, McMahon AP, Tabin CJ, Bumcrot DA, Marti-Gorostiza E;
XX WPI; 2003-799823/75.
XX DR N-PSDB; ADD25289.
XX XX Novel isolated antibody which is immunoreactive with a vertebrate
XX PT hedgehog protein sequence that binds with patched receptor, useful for
XX PT blocking action of naturally occurring hedgehog protein, and for
XX PT inhibiting spermatogenesis.
XX PS Disclosure; SEQ ID NO 62; 120pp; English.
XX CC The invention relates to an isolated antibody (I) which is immunoreactive
XX CC with a hedgehog polypeptide (II) that binds to a patched receptor, where
XX CC (II) is encoded by nucleic acid which hybridise to a fully defined
XX CC vertebrate hedgehog (hh) protein. (I) is useful as a hedgehog antagonist
XX CC by blocking action of naturally occurring hedgehog protein, and therefore
XX CC for inhibiting spermatogenesis. (I) is also useful for inhibiting normal
XX CC ovarian function. (I) is useful for blocking the action of one or more
XX CC hedgehog proteins and allows the study of the role of these proteins
XX CC e.g., embryogenesis and/or maintenance of differential tissue. (I) is
XX CC also useful in immunohistochemical staining of tissue samples in order to
XX CC evaluate the abundance and pattern of expression of the hedgehog
```

CC polypeptides. (I) is also useful diagnostically in immunoprecipitation
CC and immunoblotting to detect and evaluate hedgehog protein levels as a
CC part of clinical testing procedure. The present sequence represents the
CC amino acid sequence of a hedgehog associated polypeptide.
XX
SQ Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 7; Length 1299;
Best Local Similarity 20.4%; Pred. NO. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSSTHTLLQFGPGGCTWASNPLTILVSLVWVALAAGLVETELTDPV 380
DB 43 KKGAKSGTAIYLRVSFOSHTLGGSSVQKMGKVLVAILVLTFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHQHFG-PPFRFTNQVILTA---PNRSSRYDSLLGLPKNFSGIL 436
DB 103 QLMIQEGGGLAEALAYTKTIGEDESATHQLLIQTTHDPAASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLRLHQVSPQAQNISLQDICYAPLNPDNTSLY-----DCC 486
DB 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMPTSPFEGIYYIEQILRHLPSCS 204
QY 487 INSLIQYFQN-----NRTLLLLLTAN-----QTLMGQTSQVDMKDH 521
DB 205 IITPLDCFWEGSQLGPESAVVPIGLNQRLTWTILNPASVMQMKQKSEKISDFETV 264
QY 522 FLYCANAPITFDGTALALSCM-----ADYCAPVFPFLAIG--GYK 560
DB 265 EQMKKEAATA---SGYMERKPLNPLNCPDTPAPNKNSTQPPDVG-----ILSGCYGYA 317
QY 561 GKDYSEAEALIMTFSLNYPAGDPLQA-----589
DB 318 AKHMWPEELIVGGAKRN-RSGHLKKAQALQSVOQLMTEKEMYDQWQNYKVHILGWTOE 376
QY 590 -----KLWEEAF---LEMBRAFORMAGMFOVTTAERSLDEINRTTAEDLPFATSY 640
DB 377 KAAEVUNAKQRFNREVEQLLRQSRATNYDIYFSSNAALDDILAKSHPSALSIVIGV 436
QY 641 IVIFLYISALGSSYSWSRVVDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVILQ 700
DB 437 AVTVLY---AFCTLLRW-RDPVRQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNAATQ 492
QY 701 VDFPLVSVGADNIFIVLEYQLPRPCEPREVHIGRALGVAPSPMLCSLSEALCFPL 760
DB 493 VWFPLALGLGVDFHMLTAAYAESNR--EQTKL-----ILKKVGPSELFSACSTAGSPPA 546
QY 761 GALTMPAVRTFALTSGLAVIDFLQMSAFVALLSDSKQEAASRLDV-CCCV-----K 814
DB 547 AAFIPVPALKVFCLOAAIVMCSNLAALLVFPAMI SLDLRRRTAGRADIFCCCFPWKEQ 606
QY 815 PQBELPP-----PQGEGL-----LLGFF 832
DB 607 PKVAPPVPLNNNGRGARHPKSCNNRVALPAQNPLLEQRAIDPGSSHSLSAFLATPA 666
QY 833 QKAYAPELLHWITRGVLLFLALFVGLSYSMCHI SVGLDQELAPKDSVLLDYFLFLNR 892
DB 667 FQHYTFLMRSWKFLTMVGLAALISSLYASTRLOQGLDIDLVDPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVVTTLGYNFSSEAGNAICSSAGCNFSTQKIQYATEPPEQSYLAIP--- 949
DB 727 LF--GFYSMAVVTQ-----NFEYPTQQLLRDY-HDSPRVPHVI 764
QY 950 -----ASSWDDFDWLTSSCCRLYISGPNKDKFCPTVNSLCKNC-----993
DB 765 KNDNGGLPFDWLLIFSEWL-----GNLQKIFDEYDRGLTKECWFPNASSDA 812
QY 994 -----NSITGWSVRPSVEQ-----FKYL-PWFLNDRPNIKCPKGG 1028
DB 813 ILAYKLIVQTHGVNDVPDKELVLTNRLVNSDGLINQRAFNYLSAW-----ATWASPTTEL 868
QY 1029 LAA--YSTSVNLTSDDGVLASRFRMAYHKPLKNSQD-----YTEALRAAR 1070

DB 869 LRANCIRNRANGASQGLYPEPRQYFHP--NEYDLKIPKSLFLVYAQMFPYLHGLTDT 926
QY 1071 ELAANITADLRKVPCTDPAFEV--FPYTTNVFVEQVLTILPGLFMLSCLVPTPAVSC 1128
DB 927 QI-KTLIGHIRDLVSRYEGFGLPNYPGIPFIWEQYMT-LRSSLAMILACVLLAALVIV 984
QY 1129 LILGLDLRSGLLMLLSIVMILVDVTFGMALWDISYNAVSLINLVSAVMSVFEVSHITRS 1188
DB 985 SLLLSWAAVAVILSVLASLAQIFGAMTLLGKLSAIPAVILISVGMMLCF--NVLLIS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQLIQIFPRLNLLITL 1248
DB 1043 LGFWTSYGNRRVQLSQMSQMLGLFVHGMLTSGVAVFMLTSPPEFVIRHPCWLLLVLC 1102
QY 1249 LGLLHGLVFLPVILSVYVDPVNPALALEQKRAEAAVAAVMVASCPNHPSPRSVSTADNIYVN 1308
DB 1103 VGACNSLULVPILLUSVMGPEAE-LVPLE-----HPDRISTSPPLVR 1143
QY 1309 HS 1310
DB 1144 SS 1145
RESULT 81
ABW00878
ID ABW00878 standard; protein; 1299 AA.
XX
AC ABW00878;
XX
DT 15-JAN-2004 (first entry)
XX
DE Drosophila patched protein.
XX
KW Cell differentiation; Desert hedgehog; Dhh; Sonic hedgehog; shh;
KW Indian hedgehog; Ihh; skeletogenesis; degenerative disorder; ischaemia;
KW Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;
KW Huntington's disease; multiple sclerosis; Pick's disease; aging process;
KW trauma; anoxia; antisense gene therapy; neuroprotective; anticonvulsant;
KW neotropic.
XX
OS Drosophila sp.
XX
PN US2003186357-A1.
XX
PD 02-OCT-2003.
XX
PF 05-JUN-1995; 95US-00462386.
XX
PR 30-DEC-1993; 93US-00176427.
PR 14-DEC-1994; 94US-00356060.
PR 04-MAY-1995; 95US-00435093.
XX
PA (INGH/) INGHAM P W.
PA (MCMA/) MCMAHON A P.
PA (TABI/) TABIN C J.
XX
PI Ingham PW, McMahon AP, Tabin CJ;
XX
DR WPI; 2003-803151/75.
DR N-PSDB; A062116.
XX
PT Modulating cell growth, differentiation or survival, for treating
PT neurodegenerative diseases, such as Alzheimer's or Parkinson's disease,
PT comprises contacting the cell with a hedgehog polypeptide.
XX
PS Disclosure; Page 93-97; Opp; English.
XX
CC The present invention relates to a novel method for modulating growth, the
CC differentiation or survival of a cell. The method involves contacting
CC cell with a hedgehog polypeptide such as Desert hedgehog (Dhh), Sonic
CC hedgehog (shh) and Indian hedgehog (Ihh). The method is used to induce a
CC cell to differentiate to a neuronal cell phenotype. It is used to
CC modulate skeletogenesis. The method is used to treat a degenerative

(HARD) HARVARD COLLEGE.
Ingham PW, McMahon AP, Tabin CJ;
WPI; 2004-794735/78.
N-PSDB; ADU26448.
Modulating growth, differentiation, or survival of a mammalian cell responsive to hedgehog induction, useful for treating e.g. Alzheimer's disease, comprises contacting or treating the cell with a hedgehog polypeptide.
Disclosure; SEQ ID NO 48; 121pp; English.
This invention relates to a novel family of genes referred to as the 'hedgehog' gene family that are involved in pattern formation in embryonic cells. Specifically, it refers to genes that have a broad involvement in the formation and maintenance of ordered spatial arrangements of differentiated tissues in vertebrates. The present invention describes a method for modulating one or more of growth, differentiation or survival of a mammalian cell or an animal responsive to hedgehog induction that comprises contacting or treating the cell with a hedgehog polypeptide that mimics or antagonises the naturally occurring hedgehog protein of the cell. Furthermore, the invention provides methods for inducing a cell to differentiate into a neuronal cell phenotype and also a method of modulating skeletogenesis so as to cause one or both of chondrogenesis and osteogenesis in the target tissue. As such, pharmaceutical compositions derived thereof can be used to treat a degenerative disorder of the nervous system i.e. one that is characterised by neuronal cell death and causes, relative to the absence of hedgehog treatment, prolonged survival of neural cells in the patient. Accordingly, it can be used to treat conditions including neuromuscular disorder, autonomic disorder, central nervous system disorder, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Pick's disease, Huntington's disease and multiple sclerosis. These compositions are shown to exhibit antiparkinsonian, neuroprotective, cerebroprotective, neurotropic, anticonvulsant and vasotropic activities. This polynucleotide sequence is the Drosophila patched protein, a receptor of the hedgehog protein of the invention. NOTE: This sequence is provided in the sequence listing but is not further referred to
Sequence 1299 AA;
Query Match 6.7%; Score 464.5; DB 8; Length 1299;
Best Local Similarity 20.4%; Pred. No. 6.7e-33;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
321 KGTSLSKLSFSTHLLGQFQGWFWASWPLTILVLSVPPVLAAGLVFTLTDPV 380
43 KGKAGSRRTAIVLSRVFQSHLETGSSVQKAGKVLFAVLVSTFCVGLKSAQIHSKVH 102
381 ELWAPNSQARSEKAFHQHFG-PPRTNOVILTA---PNRSSRYVDSLLGPKFSGIL 436
103 QLMWQGGGLEALAYTKTIGEDSATHOLLITQTHDPAASVLHPQALL-----A 153
437 DLDLLELELEQLERLHLQVSPQARNISLODICVAPLNPNTSLY-----DCC 486
154 HLEVLKATAVKVLHYDTE-W-----GLRDMCMWPSTPFGIYYIEQILRHLPDS 204
487 INSLLOYFQN-----NRTLLLTAN-----QTLMGQTSQVDWQKH 521
205 IITPLDCFEGSQGLGPESAVVTPGLNQLRLLTTLNPASVMQVMQKSEEEKISFPDFTV 264
522 FLYCANAPITFDGTALALSCM-----ADYGAPVPPFLAIG--GYK 560
265 EQYMKRAAIA---SGYMEKPCINPLNPNCPDTPAPKNKSTQPPDVG-----ILSGGCGYA 317
561 GKDYSEAEALIMTFLSNLYPAGDPLAQA-----589
318 AKHMHWPPELIVGGAQRN-RSGHLKXAQALQSVVQLMTEKEMYDQWQDKYVHHLGWQGE 376
590 -----KLWEEAF---LEEMRAQRRMAGMFQVTFTEAERSLEDEINRTTAEDLPATSY 640

DB 377 KAAEVLNMQRNFSREVEQLLRQSRITATNYDIYFSSAALDDILAKFSHPALSIVIGV 436
QY 641 IVIFLVISIALGSSYSSWSVMVDSKATLGLGGVAVVVLGAVMAAMGPPSYGLRSSVLVLO 700
DB 437 ATTVLY---AFTLLRW-RDPVRGSSVGVAGVLLMCFSTAAGLGUSALLIGIVFNAASTQ 492
QY 701 VVPFLVLSGADNIFIFVLEYQYLRPRRPGEPREVHTGRALGRVAPSKMLCSLSEAICFFL 760
DB 493 VVFLALGLGVDFHIFMLTAAYAESNER--EQTKL-----ILKKVGPSILSACSTAGSPPA 546
QY 761 GALTMPAPVTRTALTSLGLAVILDPLQMSAFVALLSDSKRQBSKLDV-CCCV-----K 814
DB 547 AAAPFVPALKVFCLOAAIYMCNSLAALVFPFAMISLDLRRRTAGRADIFCCCFVPMKEQ 606
QY 815 PQELPP-----PQGGSL-----LLGPP 832
DB 607 PKVAPVPLANNNGGARHPKSCNNRVALPAQNPLLEQRADIPGSSHSLSFSLATFA 666
QY 833 QKAYAPFLHWTIRGVVLLFLALFGVSLYSCHISVGLDQELALPKDSYLDYFLFLNR 892
DB 667 FQHYTPFLMRSMVKPLTVMGFLAALISSLYASTRLQDGLDIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVTVTLGNFSSBAGMAICSSAGCNPFSFTQKIQYATEPPEQSYLAIP--- 949
DB 727 LF--GFYSMYAVTQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDPFDIMLTSSCCRLYISGNKDKFCFSTVNSLNLKNC----- 993
DB 765 KNDNGLPDFWLLLPSEWL-----GNLQKIPDEEYRDGRLTKECWFFPNASSDA 812
QY 994 -----MSITWGSVRPSVEQ-----PHKYL-PWFLNDRPNKCPKGG 1028
DB 813 ILAYKLIQVTHVDPVDKELVLTNRLVNSDGIINQAFVNYLSAW-----ATWASSPTL 868
QY 1029 LAA--YSTSVNLTSDGQVLASRFMAYHKPLKNSQD-----YTEALRAAR 1070
DB 869 LRANCIRNRANGASQOKLYPEPRQVHQD--NEYDLKIPKSLPVVAQMPFYLHGLTDS 926
QY 1071 ELAANITADLRKVPCTDPAFV--PPYTTITNVFYEQYLTILPEGLFMLSCLVPTFAVSC 1128
DB 927 QI-KTLIGHIRDLVKYEGFGLPNPSPGIPPIFWBQYMT-LRSSLAMILACVLLAALVLV 984
QY 1129 LLLGLDLRSGLLNLSIVMLVDTVGFMAWDISYNAVSLINLVASGVMSVEFVSHTRS 1188
DB 985 SLLLSVMAAVLVILSVLASLAQIFGAWTLGLKLSAIPAVILILSVGMULCF--NVLS 1042
QY 1189 FAISTKPTWLERAKATISMGSAVFAVAMTNLPGILVLGLAKAQIIOIFFRLNLLITL 1248
DB 1043 LGFMTSVGNRRQRRVQLSMQMSLGPLVHGMLTSGVAVFMLSTSPFBFVIRHFCWLLLVLC 1102
QY 1249 LGLLHGLVPLVILSVGPDVNPALALEQKRAEEAVALVAVASCPNHPSEVSTADNIYN 1308
DB 1103 VGACNSLIVFPILLSMVGPEAB-LVPLE-----HPDRISTPSPLPVR 1143
QY 1309 HS 1310
DB 1144 SS 1145
RESULT 86
ADW25903
ID ADW25903 standard; protein; 1299 AA.
XX
AC ADW25903;
XX
DT 24-MAR-2005 (first entry)
XX
DE Drosophila sp. patched protein.
XX
KW Differentiation; cell growth; pharmaceutical; patched protein.
XX
OS Drosophila sp.

Key	Location/Qualifiers	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	RESULT 87
Domain	93..426	QY	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	AE822486
Domain	700..966	QY	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	AE822486
Domain	/note = Extracellular domain	QY	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	AE822486
Domain	/note = Extracellular domain	QY	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	QY	Db	AE822486
XX	US2005004028-A1.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	06-JAN-2005.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	28-APR-2004; 2004US-00835517.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	30-DEC-1993; 93US-00176427.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	14-DEC-1994; 94US-00336060.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	04-MAY-1995; 95US-00435093.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	05-JUN-1995; 95US-00462386.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	20-OCT-1997; 97US-00954771.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	(HARD) HARVARD COLLEGE.	318	590	377	641	437	701	493	761	547	815	607	833	667	893	727	950	765	994	813	1029	869	1071	927	1129	985	1189	1043	1249	1103	1309	1144	AE822486
XX	(IMCR) IMPERIAL CANCER RES TECHNOLOGY LTD.	318	590	377	641	437	701	493	761	547																							

Query Match	6.3%	Score 438.5;	DB 2;	Length 1285;
Best Local Similarity	20.0%	Pred. No. 1.8e-30;		
Matches	242;	Conservative 174;	Mismatches 450;	Indels 347; Gaps 44;
QY	321	KGTSLSOKLSFSTHTLLGQFFQGGTWTWASPLILVLSPVVALAAGLVFTELTTDPV	380	
DB	43	KGKAGRSRTAIYRSVFQSHLETGLSSVQKHAGKYLFAVILVLSTFCVGLKSAQIHKVH	102	
QY	381	ELMSAPNSQARSEKAFHDQHEG-PPFRTNOVILTA--PNRSSYRYDSLLGLGPKNFSGIL	436	
DB	103	QLMWIOEGGRLEAEALAYTKTTIGEDSATHQLLIQTHDPNASVLHPQALL-----A	153	
QY	437	DLDLLELELEQLERLHLQVMSPEAQRNLSQDICYAPLNPDTSLY-----DCC	486	
DB	154	HLEVLVKATAVKVHLDYTE-W-----GLRDMCNMPSSTPSFEGYIEQILRHLP	204	
QY	487	INSLIQYFQN-----NRTLLLLLTAN-----OTLMGQTSQVMDKDH	521	
DB	205	IITPLDCFEGSQQLLGPESAVVIGLNQRLLTWTLNPNASVMQYMKQKMSBKISFDFETV	264	
QY	522	FLY-----C--ANAPLTFKD-----GTALALSCMADYGAAPVFP	553	
DB	265	EQYMKAAALGSGVMEKPCLPNPNCPDTPAKNKNSTQPDVGAILLSGGCYG-YAAKHMHW	323	
QY	554	---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----	590	
DB	324	PEELIVGGRKRNRSGLRKAQALQSVVQLMTEKEMVDQWQYKVVHLGWTQEKAAEVLN	383	
QY	591	LWEEAF---LEEMRPAQRMRMAGHFQVTPTAERSLEDEINRTTAEDLPFATSYVIEFYI	647	
DB	384	AWQRNFSREVEQLLRKQSRATNYDIYFSSAALDDILAKFSHPGALSIVIGVATVLY-	442	
QY	648	SLALGSYSMSRVMVDSKATLGLGVAVVLGAVMAAMGFPSYILGI-----	692	
DB	443	--AFCTLLRW-RDPVRGQSSVGAVGLMCFSTAGLGLSALLGIVFNALTAAYAESNR	499	
QY	693	-RSSLVI-----LQVPFVLVSGADNIFVLEYQRLPRRPGREPVEVHIGRALGRVAPSM	747	
DB	500	EQTLLILKNASTQVVPFLALGLVDHIFI-----VGPSTI	533	
QY	748	LLCSLSEAI CFFLGALTPMPAVTFPALTSGLAVILDFLLQMSAFVALLSLDSKQASRL	807	
DB	534	LFSACSTAGSFFAAAFIPVPAKVFCLQAAIVMCSNLAAALLVFPAMISLDRRTTAGRA	593	
QY	808	DV-CCCV-----KQELPP-----PGQ	823	
DB	594	DIFCCFPVWKQKQVAPVPLNANNRGARHPKSCNNRVPLPAQNPILLEQRADIPGS	653	
QY	824	GEGJ-----LLGFFQKAYAPFLHLWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPK	879	
DB	654	SHSLASFSLATFAFOHYTPELMSRWKFLTVMGFLAALISLYASTRLQDGLDIIDLVPK	713	
QY	880	DSYLLDFLFLNRFFEYVAPVYVTTLYGYNFSSEAGNNAICSSAGCNNTSFQKIQVATE	939	
DB	714	DSNEHFLDAQTRLF--GFYSMTAVTQ-----NEYPTQQQLLRD	752	
QY	940	FPEQSYLAIP-----ASSWVDDFIDWLTSPSSCCRLYISGPNKDKFCPSTVNSLNC	989	
DB	753	YHDS--FRVPHVKNNGGLPDFLLLFSEWL-----GNLQKIFDEYRDGRL	798	
QY	990	LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF	1015	
DB	799	TKECWFPNASSDAILAYKLI VQTGHVDNPVDKELVTLNRLNSDGIINORAFNYLSAWA	858	
QY	1016	LNDRENIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSQD-----	1061	
DB	859	TND-----VFAYG-----ASQKLIPEPQYFHP--NEYDLKIPKSLPLVYAQM	901	
QY	1062	-YTEALRAARELANITADLRKVPDGTDPAFV--FFYTITNVFYEQYLTLPBGLPMLS	1117	
DB	902	PYHLGLTDTSQI-KTLIGHIRDL SVKVEGGLNPSPGIPFIWEQYMT-LRSSLAMIL	959	
QY	1118	LCLVPTFAVSCLLGLDLRSGLINLLSIVMLVDTVGFMAWDISYNAVSLINLVSAVGM	1177	

1043 LGFMTSVGNRRVQLSNQMSLGLVHGMLTSGVAVFMLSTSPFEFVIRFCWLLLVLC 1102

1249 LGLLHGLVFLPVLISYVGPVNPALAEQKRAEAAVAVASCPNHPSPRVSTADNIYN 1308

1103 VGACNSLLVFFILLUSVVGPEAE-LVPLE-----HPDRISTPSPLPVR 1143

1309 HS 1310

1144 SS 1145

SULT 88

W72972

AAW72972 standard; protein; 1285 AA.

AAW72972;

26-JAN-1999 (first entry)

Drosophila melanogaster patched amino acid sequence.

patched gene; ptc; diagnosis; treatment; developmental disorder; cancer; healing; injured tissue; spina bifida; Wnt-1 oncogene; sperm production; gene therapy.

Drosophila melanogaster.

US5837538-A.

17-NOV-1998.

06-OCT-1995; 95US-00540406.

07-OCT-1994; 94US-00319745.

(STRD) UNIV LELAND STANFORD JUNIOR.

Scott MP, Goodrich LV, Johnson RL;

WPI, 1999-023461/02.

Nucleic acid encoding vertebrate patched protein and related transformants - used to express poly:peptide(s), useful for diagnosis and treatment of developmental disorders or cancer, and in healing of injured tissue.

Example; Col 35-42; 38pp; English.

The present invention describes vertebrate and invertebrate patched (ptc) genes. Cells containing and expressing the ptc gene are used for the recombinant production of the protein. These in turn are useful: (i) for generating antibodies (Ab); and (ii) to screen for specific-binding ligands (potential therapeutic agonists and antagonists). The ptc gene, or its fragments, are used to isolate related sequences from other mammals; to identify mutations (particularly those associated with genetic diseases such as spina bifida and other developmental disorders); to monitor expression levels in testis (to determine relationship with sperm production) and to isolate 5'-non-coding sequences (used to study embryonic development and to provide regulated expression of proteins). The complete gene can be used in gene therapy, including expression of antisense molecules, and to generate transgenic animals for studies of embryonic development. Ab are used diagnostically to determine the ptc protein on cell surfaces and as competitive inhibitors of signal transduction through the ptc ligand. Cells that have been engineered to express the ptc protein can be used to promote regrowth and healing of damaged tissue (e.g. growth of new teeth) and regulation of the ptc protein expression may be useful in cancer treatment (it may control the Wnt-1 oncogene). The present sequence represents Drosophila melanogaster (fly) patched amino acid sequence, from the present invention

Sequence 1285 AA;

```

Db 960 ACVLLAALVLSVLLLSVAAVLLVLSVLASQAIFGAMTLTGKLSAIPAVILSVGM 1019
QY 1178* SVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQIQL 1237
Db 1020 MLCP--NVLISLGFMTSVGNRRVQLSMQMSGLPLVHGMILSGVAVFMLSSTPFEVIR 1077
QY 1238 PFRLNLLTLGLLHGLVFLPVLSYVGPDPVNPALALSKRAEAAVAAVMVASCNPHS 1297
Db 1078 HFCWLLLVLCVGCACNSLLVFPILLSVWGPEAE-LVPLE-----HPD 1118
QY 1298 RYSTADNIYVNH 1310
Db 1119 RISTPSLPVRSS 1131

RESULT 89
AAB67157
ID AAB67157 standard; protein; 1285 AA.
XX
AC AAB67157;
XX
DT 12-APR-2001 (first entry)
XX
DE Fruit fly patched protein.
XX
KW Fruit fly; patched; PTC; segment polarity; limb patterning; development;
KW hedgehog; antibody; human; mouse.
XX
OS Drosophila melanogaster.
XX
XX US6172200-B1.
PN
PD
XX
XX 09-JAN-2001.
XX
XX 20-OCT-1997; 97US-00954668.
XX
XX 07-OCT-1994; 94US-00319745.
XX
XX 06-OCT-1995; 95US-00540406.
XX
XX (STRD ) UNIV LELAND S STANFORD.
XX
XX Scott MP, Goodrich LV, Johnson RL;
XX
XX WPI; 2001-136884/14.
XX
XX N-PSDB; AAF32178.
XX
XX Novel monoclonal antibody useful in diagnostic assays for detection of
PT presence of protein on surface of cells specifically binds to naturally
PT occurring patched protein, other than Drosophila patched protein.
XX
XX Disclosure; Col 35-42; 39pp; English.
XX
XX The present invention provides a monoclonal antibody which specifically
CC binds to a patched protein (PTC) other than that from Drosophila. Also
CC given are the protein and coding sequences of patched from the beetle,
CC mouse, butterfly, fruit fly, mosquito and human. Patched is a segment
CC polarity gene involved in limb patterning. The sequences can be used to
CC study development and to isolate the patched ligand, hedgehog. In
CC addition, antibodies can be used to detect the PTC protein on cell
CC surfaces or to inhibit the transduction of signal by the PTC ligand by
CC competing for its binding site
XX
XX Sequence 1285 AA;

Query Match 6.3%; Score 438.5; DB 4; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLDKLSFSTHTLLGCGFGWGTWASWPLTILVLSVIPVVALAAGLVPTETLDPV 380
Db 43 KKGKAGRTAIYLRVSFQSHLETGSSVQKAGKRVLFVAILVLSTFCVGLKSAQTHSKVH 102

```

```

QY 381 ELMSAPNSQARSEKAFHQHFG-PPFRTNQVILTA---PNRSSRYRYSLLGPKNFSGIL 436
Db 103 QLWIQEGGRLAELAVTKTIGEDSATHQLLIQTHDENASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLHLQVMSPEAQRNLSIODICVAPLNPONTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCMNPSTPSPFEGIYYIEQILRHLIPCS 204
QY 487 INSLIQYFON-----NRTLILLLTAN-----OTLQMGQTSQVDPKDH 521
Db 205 IITPLDCFWEGSQLGPESAVVPLGLNQRLLTTLNPASVMQMKMKSEKISFPFETV 264
QY 522 FLY-----C-----ANAPLTKD-----GTALASCMDYAGAPVPPF 553
Db 265 EYMKRAAIGSGYMEKPCINPLNCPDTPAPKNKSTQPPDVGAILSGGCVG-YAAKGMHW 323
QY 554 ---LAIGGYK---GKDYSEABALIMTFSL-----NNYPAGDPLAQAK----- 590
Db 324 PEELIVGGRKVRNRSGLHRAQAQLQSVVQLMTBKEMYDOMQDNYKVHHLGWTQSKAAEVLN 383
QY 591 LWEEAF---LEEMRAFORRMAGMFOVTFPAERSLEDEINRTTAEDLPIFATSVIVFLYI 647
Db 384 AWQNPFSREVEQLLRKQSRATNYDIYVPSAALDDILAKFSPSALSIVIGVAVTVLY- 442
QY 648 SLALGSYSWSRVMDSKATLGLGGVAVVLGAVMAAMGFFSYLGI----- 692
Db 443 ---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI-----LQVPPFLVLSVGDNIPIFVLEYORLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVVPLALGLGVDFHFI-----VGPSTI 533
QY 748 LLCSLSEACPFPLGALTMPAVRTEALTSLGLAVILDQLQMSAFVALLSLDSKROQASRL 807
Db 534 LFSACSTAGSPFAAFAIPVPALKVFCLOAIVMCSNLAAALVFPAMISLDRRTAGRA 593
QY 808 DV-CCCV-----KQOELPP-----PG 823
Db 594 DIFCCCFPMWKEQPKVAPVPLPLNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGS 653
QY 824 GEGL-----LLGFFQKAYAPFLHWTIRGVVLLFLALFGVLSYSMCHISVGLDQELALPK 879
Db 654 SHSLASFSLATPAQHYPFLMRSWVKFLTWGFLAALISSLYASTRLQDGLDIIDLVPK 713
QY 880 DSYLLDYLFLNRYFEVGAPVYVTTLGYNPFSSAGMNAICSSAGCNCNPFQKIQAYTE 939
Db 714 DSNEKFLDAQTRLP--GFYSWIAVTQG-----NFEYPTQQLLRD 752
QY 940 PFEQSYLAIP-----ASSWYDDFDIMLTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNGGLPDPFLLLFSEWL-----GNLQKIPDEYRDGRL 798
QY 990 LKNC-----MSITMGSVRPSVEQ-----PHKYL-PWF 1015
Db 799 TKECWFPNASSDAIILAYKLIYQTHVDNPNVDEKELVLTNRLVNSDGIINQRAFYNLSAWA 858
QY 1016 LNDRENIKCPKGLAAYSTSVNLTSDGOVLASRFMAVHKPLKNSOD----- 1061
Db 859 TND-----VPAIG-----ASQGLYPEPROYFHP--NEYDLKIPKSLPLVYAQM 901
QY 1062 --YTEALRAARELANITADLRKVPGTDPAPEV--FPYITNVFYEQVLTILPEGLFMLS 1117
Db 902 PLYLHGLTDTSQI-KTLIGHIRDLRSVKYEGFGLPNYPGIPPIFWEQYWT-LRSSLAMTL 959
QY 1118 LCLVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDISYNAVSLINLAVSGM 1177
Db 960 ACVLIAALVLSLLLSVAAVVLVLSLAQAIFGAMTLTGKLSAIPAVILSVGM 1019
QY 1178 SVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQIQL 1237
Db 1020 MLCP--NVLISLGFMTSVGNRRVQLSMQMSGLPLVHGMILSGVAVFMLSSTPFEVIR 1077
QY 1238 PFRLNLLTLGLLHGLVFLPVLSYVGPDPVNPALALSKRAEAAVAAVMVASCNPHS 1297
Db 1238 PFRLNLLTLGLLHGLVFLPVLSYVGPDPVNPALALSKRAEAAVAAVMVASCNPHS 1297

```

Db	1078	HFCMLLVLCVGCACNSILVFPILLSMVGPEAE-LVPLE-----	HPD 1118
Qy	1298	RVSTADNIYVNH5 1310	
Db	1119	RISTPSLPVRSS 1131	
RESULT 90			
AAG79574			
ID	AAG79574	standard; protein; 1285 AA.	
AC	AAG79574;		
XX	23-DEC-2002	(first entry)	
XX	Drosophila	patched protein.	
KW	Patched; ptc; basal cell nevus syndrome; BCNS; wound healing;		
KW	developmental disorder; spina bifida; craniofacial abnormality;		
KW	basal cell carcinoma; skin; brain tumour; transgenic; mouse;		
KW	proliferation; oncogenesis; signal transduction; cancer; aging.		
XX	Drosophila melanogaster.		
XX	Key	Location/Qualifiers	
FT	Misc-difference 758	/note= "Encoded by GTCCGG"	
FT	Misc-difference 812	/note= "Encoded by ATG"	
FT			
XX	US6429354-B1.		
XX	06-AUG-2002.		
XX	22-AUG-1997;	97US-00918658.	
XX	07-OCT-1994;	94US-00319745.	
PR	06-OCT-1995;	95US-00540406.	
PR	31-MAY-1996;	96US-00656055.	
XX	(STRD) UNIV LELAND STANFORD JUNIOR.		
PA	(REGC) UNIV CALIFORNIA.		
XX	Scott MP, Goodrich LV, Johnson RL, Epstein E;		
XX	WPI; 2002-722086/78.		
DR	N-PSDB; ABA00377.		
XX	Transgenic mouse, useful for cancer drug screening, comprises disruption		
PT	of allele of patched locus which predisposes the mouse to develop a		
PT	proliferative disorder.		
XX	Example; Col 57-64; 58pp; English.		
PS			
XX	This sequence shows drosophila patched (ptc) protein. Mutations in the		
CC	patched gene are characteristic in basal cell nevus syndrome (BCNS), an		
CC	inherited disorder with an increased risk of developmental disorders such		
CC	as spina bifida and craniofacial abnormalities, basal cell carcinoma of		
CC	the skin and brain tumours. The human and mouse ptc sequences show 86%		
CC	homology. The transgenic mouse of the invention has a genome comprising a		
CC	disruption of at least one allele of a patched locus, where the		
CC	disruption predisposes the mouse to develop a proliferative disorder. The		
CC	transgenic mouse is useful for studying patched function and regulation,		
CC	for e.g. a series of small deletions and/or substitutions made in the		
CC	patched gene, to determine the role of different exons in oncogenesis,		
CC	signal transduction, etc; and for drug screening which is useful for		
CC	treating cancer or developmental abnormalities attributed to a defect in		
XX	patched function such as wound healing and aging		
XX	Sequence 1285 AA;		
SQ			
Query Match			
		6.3%;	Score 438.5; DB 5; Length 1285;

Best Local Similarity 20.0%; Pred. No. 1.8e-30;			
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;			
Qy	321	KGTSLSDKLSFSTHLLGQFFGQGWTVASWELTILVLSVIPVVALAAGVFTTELTDPV 380	
Db	43	KGKARGRTAIVLRSVFQSHLETLGSSVQKHAGKLVFVAILVLTFCVGLKSAQIHSKVH 102	
Qy	381	ELWSAPNSQARSEKAFHDQHFQ-PPERTNQVILTA---PNRSYRYDSLLGPKNFSGIL 436	
Db	103	QLMIQGGRLAEALATQKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153	
Qy	437	DLDLLELELEQLERLRLQVMSPEAQRTISLDQICVAPLNPDNTSLY-----DCC 486	
Db	154	HLEVIVKATAVKVHLYDTB-W-----GLRDMCNMPSTPSPEGIYIIBQILRHLPSCS 204	
Qy	487	INSLLOVFQN-----NRTLILLTAN-----OTLMQSTQSDVDKDH 521	
Db	205	IITPLDCFWEGSQLGPESAVIPGLNQLLTLTLPASVMQMKQMSSEKISFDETV 264	
Qy	522	FLY-----C-----ANAPLTPKD-----GTALALSCMADYGAPEPF 553	
Db	265	EQMKRAAIGSGYMEKPCPLNPNCPDTAPKNKSTQPPDVGAILSGCYG-YAAKMHW 323	
Qy	554	---LAIGGYK---GKDYSEAEALIMTSL-----NNTPADGPRLAQAK----- 590	
Db	324	PEELIVGGRKRNRSGLRKAQALQSVVQLMTEKEMVDQMDNYKVHHLGWTQEKAAEVLN 383	
Qy	591	LWEEAF---LEEMRAFORRMAGMFQVTTFAERSIIDEINNTAEDLPIFATSYIVFLYI 647	
Db	384	AWQRNFRSEVEQLLRKQSIATNYDIYVFSAAUDDILAKFSPSALSIVIGNAVTVLY- 442	
Qy	648	SLALGSYSSWSRVNVDKATLGLGGVAVVLGAVMAAMGFFSYLGI----- 692	
Db	443	--AFTLLRW-RDPVRCQSSVGAGVLLMCFSTAAGLGSALLGIVENALTAAVAESNRR 499	
Qy	693	-RSSLVI-----LQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHIGALGRVAPSM 747	
Db	500	BQTKLILKNASTQVVPFLALGLGVDHIF1-----VGPSI 533	
Qy	748	LLCSLSEAI CFFLGALTPMPAVRTFALTGLAVILD FLLQMSAFVALLSLDSKRQESRL 807	
Db	534	LFSACSTAGSFPAAFIPVPALKVFCLOAIWCSNLAAALLVPAMISLDRRTAGRA 593	
Qy	808	DV-CCCV-----KPQELPP-----PGQ 823	
Db	594	DIFCCCFPWKEQPKVAPPVLPNNNGRGARHPKSCNNNRVPLPAQNPLEORADIPGS 653	
Qy	824	GEGL-----LLGFFQKAYAPFLHWHITRGVLLLLTALFGVLSYNSCHISVGLDQELALPK 879	
Db	654	SHSLASFSLATFAFOHYTFFLMRSWVKFLTYMGFLAALISSLYASTRLQDGLDIIDLVPK 713	
Qy	880	DSYLLDYFLNRYPEVGAPVVFVTTGLGYNFSSEAGMNAICSSAGCANNFSFTQKIQYATE 939	
Db	714	DSNEHKFLDAQTRLF--GFYSMYATQG-----NFEYPTQQLLRD 752	
Qy	940	FQESYLAIPE-----ASSWVDDFDIMLTPSSCCRLYISGPNKDKFCBSTVNSLNC 989	
Db	753	YHDS--FRVPHVYKNDNGLGLPDFWLLLFSEWL-----GNLQKIFDEEYRDL 798	
Qy	990	LKNC-----MSITMGSVRPSVEQ-----PHKYL-PWF 1015	
Db	799	TKECWFPNASSDAILAYKLI VQTGHVDPVDKELVLTNRLVNSDGIINQRAFNYLSAWA 858	
Qy	1016	LNDRENIKCPKGLAAYSTSVNLTSDGVLASRFMAYHKPLKNSQD----- 1061	
Db	859	TND-----VFAYG-----ASQGLKYPPEQYFHPQ--NEYDLKIPKSLPLVYAQM 901	
Qy	1062	--YTEALRAARELAANITADLRKVPGTDPAFV--PPYITNVFQYIITLPEGLFMLS 1117	
Db	902	PFLYHGLTDSQI-KTLIGHIRDLSVKYEGFGLPNYPGIPFIFWEQYMT-LRSSLAMIL 959	
Qy	1118	LCILVPTFAVSCILLGLDLRSGLNLLSIYMIIVDTVGFMAWMDISYNAVSLNLVSAVGM 1177	

Db 960 ACVLLAALVLSVLLLSVAAVILSVLASLAQIFGAMTLIGIKLSAIPAVILISVGM 1019
 QY 1178 SVEFVSHITRFAISTKPTWLERAKEATISMGSFAVAMTNLPGLVLGLAKAQLIQI 1237
 Db 1020 MLCP--NVLISLGFMTSVGNRRVQLSMQMSLGLPLVHGLTSGVAFLMLSTPPEFVIR 1077
 QY 1238 PFRLNLITLIGLLHVLPLVILSYVCPDYNPALALEQKRAEAAVAVMVASCPNHS 1297
 Db 1078 HFCWLLLVLCVGCNLSLVPFILLSMVGPEAE-LVPLE-----HPD 1118
 QY 1298 RVSTADNIYVNH 1310
 Db 1119 RISTSPPLVRSS 1131

RESULT 91

ABU62150
 ID ABU62150 standard; protein; 1285 AA.

AC ABU62150;

XX 25-AUG-2003 (first entry)

XX Fruit fly patched gene PTC product.

DE Fruit fly; patched gene; PTC; hedgehog protein; gene therapy.

XX Drosophila melanogaster.

XX US200302085-A1.

XX 13-FEB-2003.

XX 20-OCT-1997; 97US-00954701.

XX 07-OCT-1994; 94US-00319745.

PR 06-OCT-1995; 95US-00540406.

XX (SCOT/) SCOTT M P.

PA (GOOD/) GOODRICH L V.

PA (JOHN/) JOHNSON R L.

XX Scott MP, Goodrich LV, Johnson RL;

PI WPI; 2003-492065/46.

XX New DNA sequence other than present in a chromosome encoding patched gene
 PT other than Drosophila patched gene, useful for preparing transgenic
 PT laboratory animals and to knock out patched protein in embryonic stem
 PT cells.

PS Disclosure; Page 8-10; 40pp; English.

XX The invention relates to a DNA sequence other than present in chromosome
 CC encoding a patched (ptc) gene other than Drosophila patched gene. A cell
 CC expressing the DNA sequence is useful for producing patched protein. A cell
 CC growing the cell expressing the DNA sequence, where the patched protein
 CC is expressed and isolating the patched protein free of other proteins.
 CC The cell expressing the DNA sequence is also useful for screening
 CC candidate compounds for binding affinity to the patched protein, by
 CC combining the candidate protein with the cell expressing the DNA sequence
 CC where the DNA sequence comprises the entire coding sequence under the
 CC transcriptional regulation of the transcriptional initiation region and a
 CC transcriptional termination region functional in the cell, expressing the
 CC patched protein in the cell and assaying for the binding of the candidate
 CC compound to the patched protein. The above cell is useful for screening
 CC candidate compounds for agonist activity with the patched protein. The
 CC DNA sequence is useful for producing all or portions of the patched
 CC protein, as probes for research, diagnosis, binding of hedgehog protein
 CC for its isolation and purification and in gene therapy. The DNA sequence
 CC is also useful as primers for investigating other species and for
 CC isolating genes from various mammalian sources of interest, particularly
 CC from humans or from domestic animals. The DNA sequence is further useful

CC for preparing transgenic laboratory animals and to knock out the PTC
 CC protein in the embryonic stem cells, so as to produce hosts with single
 CC functional patched gene. The present sequence represents the amino acid
 CC sequence of the fruit fly patched gene PTC product
 XX
 SQ Sequence 1285 AA;

Query Match 6.3%; Score 438.5; DB 7; Length 1285;

Best Local Similarity 20.0%; Pred. No. 1.8e-30;

Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

QY 321 KGTSLSDKLSFSTHTLLGQFFQCGMTWASWPELTLLVLSVIPVVALAAGLVFELTDPV 380

Db 43 KKGARGRTAIYLRVSFQSHLETGSSVQKHAGKLVFAILLVLTSTCVGLSKAQIHSKVH 102

QY 381 ELWSAPNSQARSEKAFHQHFG-PPFRTNOVILTA---PNRSSYRVDSLLGPKNPSGIL 436

Db 103 QLMIOEGGRLEALAYTKTIGEDSATHQLIQTTHDPNASVLHPQALL-----A 153

QY 437 DIDLLLELLELQERLRLHQLVNSPEQRNLSLODICVAPLNPNTISLY-----DOC 486

Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCNMPSPTSPGIIYIEQILRHLPIS 204

QY 487 INSLLQYFQN-----NRTLLLLTAN-----QTLMGQTSQVDWKDH 521

Db 205 IITPLDCFWEGSQLLGPSAVVPIGLNORLLTTLNPASVMQYMKQKSEKISDFETV 264

QY 522 PLY-----C-ANAPLTFKD-----CTALALSCWADYGAPVFPF 553

Db 265 EGYMKRAAIGSGYMEKPCNLNPNCPDTPAPKNSTQPPDVGAIISSGGCYG-YAAKHMEW 323

QY 554 ---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590

Db 324 PEELIVGGRKRNRSGLHKAQALQSVQLMTEKMYDQODNYKVHHLGWTQEKAAEVLN 383

QY 591 LWEEAF---LEMRAPQRMAGMFVFTTAERSLSDEINRTAEDLPFATSYVIFLYI 647

Db 384 AWQRNFSREVEQLLEKQSRATNYDIYVFPSSAALDDILAKFSPHSALSIVIGVAVTVLY- 442

QY 648 SIALGSSYSWSRVMYDSKATLGLGVAVVLGVAAMGPFPSVYLG----- 692

Db 443 --APCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGVFNALTAAYAESNR 499

QY 693 -RSSLVI---LQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747

Db 500 EQTKLILKNASTQVVPFLALGLGVDFI-----VGPSI 533

QY 748 LICSISEAICPFLGALTMPAVRTFALTSGLAVIDFLQMSAFVALISLDSKROBSRL 807

Db 534 LFSACSTAGSPFAAAFIQVPAIKVFLQAAIWCNLAALVFPAMI SLDLRRRTAGRA 593

QY 808 DV-CCCV-----KQQLPP-----PQ 823

Db 594 DIFCCCFPVWKQPKVAPFVPLNNNGRGARHPKSCNNRVLPAPQNPLLBQRADIPGS 653

QY 824 GEGL-----LLGFFQKAYAFPLHMTIRGVLLFLALPGVLSYSMCHTSVGLDQSLALPK 879

Db 654 SHSLASFSLATFAFQHYTFPLMRSMVKVLTVMGFLAALISSLYASTRLQDGLDIDLVPK 713

QY 880 DSYLLDYFLNRYPEVGAPVYVTTGLNFSSEAGMNAICSSACGNFSPFTQKIYATE 939

Db 714 DSNHKKFLDAQTRLF--GFYSMYAVTQ-----NFEYPTQQLLRD 752

QY 940 FPEQSYLAIP-----ASSWDDIDMLTPSSCCCLYISGPNKDKFCFSTVNSLNC 989

Db 753 YHDS--FRVPHVKNNDNGGLPDFWLLLFSEWL-----GNLQKIFDEBYRDRGL 798

QY 990 LKNC-----MSITMGSVRPSVQ-----FHKYL-PWF 1015

Db 799 TKECWFPNASSDAILAYKLIQVTHVDNPNVDEKELVLTRELNSDGIINORAFNYLSAWA 858

QY 1016 LNDPRNPKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPLKNSOD----- 1061

Db	859	TND	-----VFAYG-----ASQKLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQM	901
Qy	1062	--YTEALRAARELAANTADLRKVPGTDPAFEV--PPYTTITNVFYEQYLITLPEGLFMLS	1117	
Db	902	PFYUHGHTDSQI-KTLGHIRDLSVKYEGFGLPNTYPSGIPPFIEWEQYMT-LRSSLAMIL	959	
Qy	1118	LCLVPTFAVSCLLGLDRLSGLLMLLSIVMLVDVTGCFMALWIDSINYANSLINLSAVGM	1177	
Db	960	ACVLLAALVLSVLSLLSVMVAVLVSATLASLAQIFGAMTLLGKUSAPAVILLIUSVGM	1019	
Qy	1178	SVEPVSHITSPASTKPTMLERAKEATISNGSAVFAGVAMTNLPGVILVLGAKAQILQI	1237	
Db	1020	MLCF--NVLLISLGFMTSVGNRRQVRVQLSMQMSLGPLVHGMLTSGVAFVMLSTSPFEFVIR	1077	
Qy	1238	FFFLNLLITLLGLHGLVFLPVTLISYVGDPVNPALALEQKRAEEAAVAAMVASCPNHP	1297	
Db	1078	HFCWLLIIVLCVGCACNSLUVFPIILLSWVGPEAB-LVPLE-----HPD	1118	
Qy	1298	RVSTADNIYVNH	1310	
Db	1119	RISTPSPLPVRSS	1131	

RESULT 92

ADP
ADE94211
ID ADE94211 standard: protein: 1285 AA.

ADE94211:

12-FEB-2004 (first entry)

fruit flv ptc polypeptide.

Fruit fly: patched gene: ptc: developmental abnormality;

FRIT1L1; patched gene; ptc; developmental disorder; proliferative disorder; tumour; oncogenic patched mutation;

proliferative disorder; tumour; oncogenic pathway mutation; patchy-associated phenotype; basal cell nevus syndrome; BCNS; carcinoma; meningioma; fibroma; cancer.

Drosophila melanogaster.

US6551782-B1.

22-APR-2003.

28-NOV-2000: 2000US-00724631.

07-OCT-1994: 94US-00319745.

06-OCT-1995; 95US-00540406.

31-MAY-1996; 96US-00656055.

22-AUG-1997; 97US-00918658.

(CTPD) INTV STANFORD

(SIRD) UNIV STANFORD.
(REGC) UNIV CALIFORNIA.

scott MP, Goodrich LV, Johnson RL, Epstein E;

WPI: 2003-615308/58.

Assay for phenotyping patched status of cell in mammalian (preferably human) cell sample, involves detecting presence or absence of aberrant modification or mutation of a patched gene, and mis-expression of patched gene.

Disclosure: SEO ID NO 6: 57pp; English.

The invention relates to an assay for phenotyping the patched status of a cell, involving detecting *in vitro*, in a sample of mammalian cells, the presence or absence of a genetic lesion in a patched gene (pcr) characterised by at least one of an aberrant modification or mutation of a patched gene or mis-expression of the patched gene. The invention also relates to a method for diagnosing a genetic predisposition of an animal to a developmental abnormality or a proliferative disorder, where the developmental abnormality or proliferative disorder correlates with

aberrant expression or activity of a patched gene or gene product, involving detecting in vitro the presence of a predisposing mutation in a patched gene in cells of the animal. Characterising the phenotype of a tumour, involves detecting the presence of an oncogenic patched mutation in cells of the tumour, where the presence of the oncogenic mutation indicates that the tumour has a patched-associated phenotype. The assay is useful for phenotyping the patched status of the cell in a mammalian cell sample obtained from a human patient. The disorders that can be detected include basal cell nevus syndrome (BCNS), carcinomas, meningiomas and fibromas. This sequence represents a fruit fly *ptc* polypeptide of the invention.

XX Sequence 1285 AA;

Query Match	6.3%	Score 438.5;	DB 7;	Length 1285;
Best Local Similarity	20.0%	Pred. No. 1.8e-30;		
Matches	242:	Conservative	174;	Mismatches 450;
		Indels	347;	Gaps 44;

QY	321	KGTSLSKLSFSTHTLQCFQFGHGTWVASWPLTILVLVSIVPVVAAGLVFTLTTDPV	380
DB	43	KGKARGSRATYLRVSFQSHLETJGSSVQKHAGKVLFAVLIVJSTFCVGLKSAIGSKVH	102
QY	381	ELWSPNQSARSEKAFHDQHFQ-PFFRTNOVILTA---PNRSSRYRDSLLIGPNKFSGIL	436
DB	103	QLWTQEGRLAEALAYTKTIGEDESATHQLLIQTHDPNASVLHPQALL---	153
QY	437	DLDLLELLELQERLRLQVWSPQAQRNISLODI CYAPLNPDPNTSLY---	486
DB	154	HLEVLVKATAVKVHLXYDE-W-----GLRDMCNMPSFSGEYIYIEQILRHLIPCS	204
QY	487	INSLLOTFQON-----NRTLILLLTAN-----QTLWGOTQSVDWKDH	521
DB	205	IITPLDCFWEGSQLGPESAVVIEGLMQRLWTTLNPASVMQYMKQKMSSEKISFDFETV	266
QY	522	FLY-----C-----ANAPLTFKD-----GTALALSCMADYGAPVFPF	553
DB	265	EQYMKRAALGSGYMEKPCLPINPCPDPTAPKNSTQPPDVGAILSGGCYG-YAAKHMHW	323
QY	554	---LAIGYK---GKDYSEAEALINTPSL-----NNYPAGDPRLLQAK-----	590
DB	324	PEELIVGCRKRNRSGLRKAQALQSVMQTEKEMYQWQDNQYKVHHIGWTQEKAAEVLN	383
QY	591	LWEEAF---LEEMRAFORMAGMFVOTFAERSLEDEINRTTAEDLPIFATSYIVIFLYI	647
DB	384	AMQRFNSREVEQLLKQKRIATNYDIYVSSNALDDIILAKSPHSALSYIVIGVATVLY-	442
QY	648	SLALGSYSWSRVMVDSKATLGLGVAVVLGAVMAAMGFFSYLGI-----	692
DB	443	--AFCTLLRW-RDPVRGQSSGVAGVLLMCFSTAAGLGLSALLGIVFNALTAAAYAESNR	499
QY	693	--RSSLVI---IQVVPFLVLSVGADNIIFVLEYQRLPRRPGCEPREVHIGRALGRVAPSM	747
DB	500	EQTKLILKNASTQVVPFLALGIVGDHIF-----VGPBI	533
QY	748	LLCSLSEATCFGLGALTMPAPVTRTALTSGLAVILDFLLQMSAFVALLSLDKRQESRL	807
DB	534	LFSACSTAGSFAAFAFIPVPAKLVFCLOARIVMCSNLAAALLVFPAMISLDLRRRTAGRA	593
QY	808	DV-CCCV-----KPOELPP-----PGQ	823
DB	594	DIFCCCFPVWKEQPKVAPPVPLPLNNNGRGARHPKSCNNNRVPLPAQNPLLEQORADIPGS	653
QY	824	GEGL-----LLGFFQKAYAPFLHLWITRGVVVLLFLALFGVSYLSYMCCHSVGLDQELALPK	879
DB	654	SHSLASFSLSATFAQHYTFPLMRWSVKVFTVWGFLLAALISLSSYASTRLQDGLDIIDLVPK	713
QY	880	DSYLLDYFLFLNRYFEVGAPVVFVTTLTYGYNFSFSEAGMNAICSSAGCNFTSFTOKI QVATE	939
DB	714	DSNEHFLDAQTELF--GFSYMAVATQ-----NFEYPTQQQLLRD	752
QY	940	FPEQSYLAIP-----ASSWDDFDIWLTPSSCCRILYISGPNKDKFCPSTVNSLNC	989
DB	753	YHDS--FRYPHVHITKNDNGGLDPLFWLLLFSEWL-----GNLQKIDDEYDRGLR	798

QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TRECFPNASSDAILAYKLI VQTGHVDNEVDKELVLTNRLVNSDGLIINQRAFYVLSAWA 858
QY 1016 LNDNRNICKCPKGLAAYSTSVNLTSDGVVLSRPMAYHKLKNSQD----- 1061
Db 859 TND-----VFAYG-----ASQGLYPPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901
QY 1062 --YTEALRAARELAANITADLRKVGCTDPAPFV--PPYITVNFVEOYLTILPEGLFMLS 1117
Db 902 PFLHGLTDTSQI-KTLIGHIRDLSVKYEGFGLPNVPSGIPFIFWEQWT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLDRSGNLNLSIVMILVDTVGFMALWDISYNAVSLINLSAVGM 1177
Db 960 ACVLLAALVLSLLLSVMAAVLIVLSLASLAQIFGAMTLGKLSAIPAVILISVGM 1019
QY 1178 SVEFVSHITRSFAISTKPTWLRKAETISMSAVFAGVAMTNLPGILVGLAKAQLQI 1237
Db 1020 MLCF--NVLLISLGFMTSVGNRRVQLSMQMSLGPLVHGMLTSGVAVFMLSTSPFEVIR 1077
QY 1238 PFERLNLITLLGLHGLVFLPVILSYGPDVNPALALBOKRABEVAAMVASCNPHS 1297
Db 1078 HECWLLVVLVGCANSLVFFILLISWVGPEAE-LVPLE-----HPD 1118
QY 1298 RYSTADNIYVNS 1310
Db 1119 RISTPSLPVRS 1131

RESULT 93

ID ADH62718
ADH62718 standard; protein; 1285 AA.

XX AC ADH62718;
XX DT 25-MAR-2004 (first entry)

XX DE Fly patched (PTC) homologue protein.
XX KW PTC; patched; embryonic development; cellular regulation;
XX KW signal transduction; fly.

XX OS Drosophila melanogaster.
XX PN US2003148388-A1.
XX PD 07-AUG-2003.

XX PF 03-JAN-2001; 2001US-00754032.
XX PR 07-OCT-1994; 9AUS-00319745.
XX PR 06-OCT-1995; 95US-00540406.
XX PR 20-OCT-1997; 97US-00954669.

XX PA (STRD) UNIV STANFORD.
XX PI Scott MP, Goodrich LV, Johnson RL;
XX DR WPI; 2003-897566/82.

XX New DNA sequences encoding patched protein from organisms including human
PT and mouse are used to elucidate embryonic development and cellular
PT regulation associated with patched signal transduction and identify
PT patched ligands.
XX Disclosure; SEQ ID NO 6; 40pp; English.
XX The present invention relates to a DNA sequence encoding a patched (PTC)
CC gene. The invention is used to elucidate embryonic development, cellular
CC regulation associated with signal transduction by the patched gene,
CC identify agonists and antagonists of that signal transduction, identify
CC and isolate ligands of patched protein and assay for levels of

CC transcription of the patched gene. The present sequence is fly patched
CC protein.
XX
SQ Sequence 1285 AA;
Query Match 6.3%; Score 438.5; DB 7; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
QY 321 KGTSLSDKLSFSTHTLLGQFFQGCWCTWASWPLTLLVLSVIPVVALAAGLVFTELTDPV 380
Db 43 KKGAGSTAIYLSRVFQSHLETGSSVQKRGKGVFLVAILVLSFVCGLKSAQIHSKVH 102
QY 381 ELMSAPNSQARSEKAFHDQHFQ-PPFRTNQVILTA---PNRSSRYDSLLGLPKNFGSL 436
Db 103 QLMIQEGGRLEAEIAYTKTIGEDESATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLRLHQLQVSPQARNISLQDICYAPLNPNSTSLY-----DCC 486
Db 154 HLEVILKATAVKVHLDTSE-W-----GLRDMCMFSTPSEFGIYIETILRHLIPCS 204
QY 487 INSLIQYFQN-----NRTLALLLTAN-----QTLMGQTSQVQDWDKH 521
Db 205 IITPLDCFWESQLLGPESAVVIPGLNQLRLMTTLNPASVMQMKQKSEKISDFETV 264
QY 522 FLY-----C-----ANAPLTFKD-----GTALALSQADYDGPVFPF 553
Db 265 EQYMKRAAIGSYMEKPCINPLNPNCPDTPAKNSTQPPDVGAILSGCYG-YAAKHMHW 323
QY 554 ---LAIGYK---GNDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db 324 PEELTVGGRKRNRSGLRKAQALQSVQLMTEKEMYDQWQDNKYVHHLGWTQEKAAEVLN 383
QY 591 LWEEAF---LSEMRAPORMAGMFOVTFTEAESLSEDEINRTTAEDLPFATSYVIFLYI 647
Db 384 AWQNFSEVSQLRKQSIATNYDIYVFSAALDDILAKFSHPALSISIVGAVTVLY- 442
QY 648 SLALGSYSSWSRVMDSKATIGLGVAVVLGAVMAAMGFFSYLGI----- 692
Db 443 --AFCTLRW-RDPVRGQSSVGAGVLLMCFSTAAGLGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI---LQVVPFLVLSVGDNIPIFVLEYORLPRRPGEPREVIHGRALGRVAPSM 747
Db 500 EOTKLILKNASTQVVPFLALGLGVDFIFI-----VGPSI 533
QY 748 LLCSLSEACFPLGALTTPBAVRTALTSLGLAVILDFLLQMSAFVALLSLDSKQAEASL 807
Db 534 LFSACSTAGSPFAAAFIIPVPAIKVFCLOAAIVMCSNLAAALLVFPAMISLDLRRRTAGRA 593
QY 808 DV-CCCV-----KQOELPP-----PGQ 823
Db 594 DIFCCCFPVWKEQPKVAPVLPVPLNNGNGRGARHPKSCNNRVLPVLPQNPPLLRADIPGS 653
QY 824 GEGE-----LLGFFQKAYAPFLHMTIRGVLLFLALFGLSVLSYMSCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFOHYTFPLMRSMVKFLTWGFLAALISSLYASTLQDGLDIDLVPK 713
QY 880 DSYLLDYFLPLNRYPEVGAPVYFVTTLGYNFSSSEAGMAICSSACNNFSPFKQIYATE 939
Db 714 DSNHFKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQOQLLRD 752
QY 940 PPEQSYLAIP-----ASSWVDDFDIMLTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNKNDGGLPDPFWLLLFSEWL-----GNLQKIPEEYRDGEL 798
QY 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLI VQTGHVDNPDVKELVLTNRLVNSDGLIINQRAFYVLSAWA 858
QY 1016 LNDNRNICKCPKGLAAYSTSVNLTSDGVVLSRPMAYHKLKNSQD----- 1061
Db 859 TND-----VFAYG-----ASQGLYPPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901

QY 1062 --YTEALRAARELANITADLRKVPGTDPDAFEV--FPYTTNNVYQYVLTILPGLPMLS 1117
Db 902 PFYHLGLTDTQI-KTLIGHIRDLDSVKYEGFGLPNYPSPGIPFIWEQYMT-LRSSLAMIL 959
QY 1118 LCLVPTFAVSCLLGLCLDRGLNLLSTVMTLVDTGFMALWDISYNVAVSLINLVSAVGM 1177
Db 960 ACVLAALVLSLLLSVMAAVLVLSVLAQIFGNTLGLKLSAIPAVILSVGM 1019
QY 1178 SVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAVAGVAMTNLPGILVLGLAKAQIUI 1237
Db 1020 MLCF--NVLSISGFMTSVGNRRVQLSMQNSGLPLVHGMLTSGVAVFMLSTSPFVIR 1077
QY 1238 PFRRLNLITLGLLHGLVFLPVILSVYGPDPVPALALEQKRAEAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVACNSLLVPFILLSMVGPEAE-LVPLE-----HPD 1118
QY 1298 RUSTADNIYVNH 1310
Db 1119 RISTSPUPVRSS 1131

RESULT 94
ADE48976
ID ADE48976 standard; protein; 1285 AA.

AC ADE48976;
XX
XX
XX 29-JAN-2004 (first entry)
XX
XX Drosophila patched protein (ptc).
XX
XX cytostatic; vulnery; gene therapy; phenotyping; patched status;
KW patched gene; genetic predisposition; basal cell nevus syndrome; tumour;
KW carcinoma; meningioma; medulloma; fibroma; cancer; wound healing; aging;
KW fruit fly; patched gene; ptc.
XX
OS Drosophila melanogaster.
XX
XX US2003186309-A1.
XX
XX 02-OCT-2003.
XX
XX 22-APR-2003; 2003US-00421446.
XX
XX 07-OCT-1994; 94US-00319745.
XX 06-OCT-1995; 95US-00540406.
XX 31-MAY-1996; 96US-00656055.
XX 22-AUG-1997; 97US-00918658.
XX 28-NOV-2000; 2000US-00724631.
XX
XX (STRD) UNIV LELAND STANFORD JUNIOR.
XX (REGC) UNIV CALIFORNIA.

XX Scott MP, Goodrich LV, Johnson RL, Epstein E;
XX
XX WPI; 2004-041193/04.
XX N-PSDB; ADE48975.
XX
XX Phenotyping the patched status of a cell for diagnosing a genetic
XX predisposition for a tumor comprises detecting the presence or absence of
XX a genetic lesion having aberrant modification, mutation or mis-expression
XX of the patched gene.
XX
XX Disclosure; SEQ ID NO 6; 60pp; English.

XX
XX The invention describes an assay for phenotyping the patched status of a
XX cell comprising detecting in a sample of mammalian cells the presence or
XX absence of a genetic lesion having aberrant modification or mutation of a
XX patched gene or mis-expression of the patched gene. The assay is useful
XX for diagnosing a genetic predisposition of an animal, e.g. basal cell
XX nevus syndrome, predisposition for developing tumour, i.e. carcinoma,
XX meningioma, medulloma, or fibroma. A genetic construct encoding a patched
XX polypeptide is used to treat an animal having a disorder comprising loss

CC of function of a wild-type patched gene, such as cancer, and can enhance
CC patch function in e.g. wound healing and aging. This sequence is encoded
CC by the fruit fly patched gene.

XX
SQ Sequence 1285 AA;
Query Match 6.3%; Score 438.5; DB 8; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.8e-30;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
QY 321 KGTSLSDKLSFSTHTLGGQFQGWGTWASWPTILVLSVIPVVALAAGLVFTELTTDPV 380
Db 43 KKGARSGRTAIYLSVFSQSHLETGLSGSVQKAGKLVFVAILVLTSTFCVQKSAQIHSKVH 102
QY 381 ELMSAPNSQARSKAFHQHFG-PFFRTNOVILTA---PNRSSRYDLSLLGPKNFGSIL 436
Db 103 QLMIQEGGRLEAEALVTKTIGEDSATHQLLQTHDPNASVLHPQALL-----A 153
QY 437 DLQLELLELQBLRLHQLVQWSPQAQRNLSLODICYAPLNPDMNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRMCNMPSTSPGIIYIEQLRLHLIPCS 204
QY 487 INSLLYQFQN-----NRTLLLTAN-----QTLMGQTSQVDMKDH 521
Db 205 IITPLDCFWEGSQLLGPESAVVIFGLNORLLTTLNPASVMQYMKQKSEKISFDFTV 264
QY 522 FLY-----C-----ANAPLTFKD-----GTALALSCMADYCAPVFPF 553
Db 265 EQYMKRAAIGSYWEKPCPLNPNCPDTPAKNKSTQPPDVGAILSGCYG-YAAKMHW 323
QY 554 ---LAIIGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAQ----- 590
Db 324 PEELIVGRKENRSGHLRKAQALQSVVQLMTEKEMYQWQDNKYVHLLGWTQEKAAEVLN 383
QY 591 LWEEAF---LEEMRAFQRMAGMEQVITFAERSLEDEINRTAEDLPIFATSIYVIFYI 647
Db 384 AWQNFREVBQLLRKQSRITATNDIYVFSAAADDILAKFHSFALSIVIGVAVTDLY- 442
QY 648 SLALGSYSSRSRVMVDKATLGLGVAVVLGAVMAAMGFFSYLGI----- 692
Db 443 --AFTLLRW-RDPVGGQSVGVAGVLLMCFSTAAGLSALLGIVFNALTAAYAESNR 499
QY 693 -RSSLVI---LQVVPFLVLSVGADNIPIFVLEYQRLPRRGEPREVHIGALGRVASM 747
Db 500 EQTKLILKNASTQVVPFLALGLGVDHIFI-----VGPSI 533
QY 748 LLCSLSEACFELGALTMPAPVTFALTSLGLAVILDFLQMSAFVALLSLDSKQASRL 807
Db 534 LFSACSTAGSFFAAFIPVPALKVFCLOAAIWMCSNLAALLVFPAMISLDRRTAGRA 593
QY 808 DV-CCCV-----KQOELPP-----PGQ 823
Db 594 DIFCCCFPWKEQPKVAPVPLPLNNNGRGARHPKSCNNRVPLPAQNPLLEQRADI 653
QY 824 GEGE-----LLGPFQKAYAPFLLHWTIRGVVLLFLALFGVLSYMSCHLSVGLDQLALPK 879
Db 654 SHSLASFSLATPAFQHYTPFLMRSVWKFLTMVGFLAALISLSYASTRLQDGLDIIDLVPK 713
QY 880 DSYLDYDFLNRYPFVGAPVVFVTLGLVFNFSSEAGMNAICSSAGCNPFSTFKIQYATE 939
Db 714 DSENHFLDAQTRLF--GFYSNAYVTOG-----NFEYPTQQLLRD 752
QY 940 FPEQSYLAIP-----ASSWVDDDFDWMITPSSCCRLYISGNKOKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVIKNDNGGLPDLFLLFSEWL-----GNLQKIFDEYRDGRL 798
QY 990 LKNC-----MSITWGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAIILAYKLIQVQTHVDNPDVKELVITNRLVNSDGLINORAFNYLSANA 858
QY 1016 LNDPRNPKCPKGLAAYSVNLTSQVQLASRFMAYHKPKNSQD----- 1061
Db 859 TND-----VFAYG-----ASQGLYPEPROYFHQP--NEYDKLPKSLPLVYAQM 901

QY 1062 --YTEALRAARELAANITADLRKVGCTDPAFEV--FPYTTITNVFEQYLTILPEGLFMLS 1117
 Db 902" PFYTLGLTDSQI-KTLIGHIRDLSVKYEGFLENYPSGIPFIWEQYMT-LRSSLAMTL 959
 QY 1118 LCLVPTFAVSCLLGLDLSGLNLLSYVILVDTVGFMAWDISYNAVSLINLSAVGM 1177
 Db 960 ACVLLAALVLSGLLSVMAVILVLSLASIAQIFGANTLLGKLSAIPAVILLSVGM 1019
 QY 1178 SVEFVSHITRFAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQI 1237
 Db 1020 MLCF--NVLLISLGFMTSGNRQRRVQLSNQMSLGPLVHGMLTSGVAVFMLSPPFVIR 1077
 QY 1238 PFRRLNLLTLGLLHGLVFLPVILSYGPDVNPALALOKRAEAAVAVMVASCPNHS 1297
 Db 1078 HFCMLLVLCVGCACNSLIVFILLISVWGPEAB-LVPLE-----HFD 1118
 QY 1298 RVSTADNIVNHS 1310
 Db 1119 RISTPSLPVRSS 1131

RESULT 95

ADN22757

ID ADN22757 standard; protein; 714 AA.

AC

XX ADN22757;

DT 02-DEC-2004 (first entry)

XX Bacterial polypeptide #5410.

DE

KW Recombinant DNA construct; transformed plant; improved plant property;

KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;

KW pathogen tolerance; pest tolerance; plant disease resistance;

KW cell cycle pathway modification; plant growth regulator;

KW homologous recombination; seed oil yield; protein yield; carbohydrate;

KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;

KW bacterial polypeptide.

XX Bacteria.

XX US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOY/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;

XX WPI; 2004-061375/06.

XX New recombinant DNA construct comprising a promoter positioned to provide

XX for expression of a polynucleotide encoding a polypeptide from a

XX microbial source, useful for producing plants with improved properties.

XX Claim 1; SEQ ID NO 5410; 122pp; English.

XX The invention relates to a recombinant DNA construct comprising a

XX promoter functional in a plant cell, where the promoter is positioned to

XX provide for expression of a polynucleotide encoding a polypeptide from a

XX microbial source. The invention also relates to a transformed plant

XX comprising the recombinant DNA construct and a method of producing a

XX transformed plant having an improved property. The plant is a crop plant

XX such as maize or soybean. The method of producing a transformed plant

CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.

XX Sequence 714 AA;

Query Match 5.5%; Score 379; DB 8; Length 714;

Best Local Similarity 23.0%; Pred. No. 2.5e-25;

Matches 173; Conservative 119; Mismatches 299; Indels 160; Gaps 25;

QY 545 DYGAFFPFLAIGGYKGDYSEAEALIMTFSLN--NYP-----AGDPRLAQ--AKLWEEAF 596
 Db 58 DYAAEM-----TGGCSG-----FASNVLNWPEDMILGNPRAKGGKLSGADA 100
 QY 597 LEEMRAFORRMAGMFQVTFRTAERSLDEINRTAEDLPFPATSYIVIFLY----- 646
 Db 101 LQ-----FVFLVASPADVFLRFKQSLSEB-----TLETLSLYERHCILHHRHPLLICSRNSV 151
 QY 647 ---ISLALGSVSSWSRVVV-----DSKATLGLGCVAVVILGAVMAAGPFPSYL 690
 Db 152 NSTPSSLLDMLSCWMLLLILLDTTAFRLQHWGFGALAGLVVTFASVAGIGLATWTF 211
 QY 691 GIRSLVILQVVPFLVLSVGADNIFIVLEYQLRPRRPFEPREPHVIGALGRVAPSMILJC 750
 Db 212 GLEFNAATTQIVPFLTIGVDNMPLLNHYRDVVKLAGGHAEMAI--LMRETGHSILCT 269
 QY 751 SLSEAI CFFL GALT PMPAVRTFALTSLGLAVILD FLLQMSAFVALLSLSKQREARLDVC 810
 Db 270 SINNLISFLTGTLLPILPALSFCQSSILLTFNFIALTITPAIISIDLRKKAKORRDF 329
 QY 811 CCVK--PQELPPPGQEGE-----LLGEFQKAVAPFLHWTIRGVLLLL 854
 Db 330 TREESTSEASIMQDFGITQOMASSDDPAPWSLHSFIRYIYFIPISKDPAKVAIIVGCC 389
 QY 855 ALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPAPVYFVTTILGYNFS--- 911
 Db 390 ALLGASFVGMRSQSTLGLSLGDLVPEHTAPAQFLRARDKYFSP-YPM-FAVIKGPNIIDYAH 447
 QY 912 SEAGNNAICSSAGCNFNFTOKIQVATEFPQSYLAIPASSWDDFIDWLTPSSCCRLYI 971
 Db 448 QQRQIDNTRQSIGSSKYVKNK---NEEPSEKY-----WGLMRDWLI--SIQRGFD 494
 QY 972 SGPNKDKP-----CPSTVNSLNCNKMSTTMSGVRSVPSEQFHK--YLPWFANDRPNI 1022
 Db 495 BEVAKGSFNLSTGTVIGNSVEDARLAHALMCSHGLFECAGIRIKIRLIPPLYT----- 549
 QY 1023 KCPKGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRK 1082
 Db 550 -----GLTDTAVIVDAIKDIRSVCFP----- 571
 QY 1083 VPGTDAPEVPYPTITNVFYQYLTILPEGLFMLSCLVPTFVAVSCLLLGLDLSGLLNL 1142
 Db 572 ---TDQGLPNFPQGIPTFWEQYLFITGNLMAQISIIITISVFCVLSVLL-FNFWAALMVV 627
 QY 1143 LSIWMLVDTVGFMAWDISYNAVSLINLSVAGVMSVEFVSHITFSPAISTKPTWLERAK 1202
 Db 628 CILGIMCELAFGMGLVGIKLNPNVSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
 QY 1203 BATISMGSAVPAVG---AMTNLPGILVLGLA 1230

FT Modified-site 192..195 /note="N-glycosylation site"
 FT Modified-site 194..196 /note="Protein kinase C phosphorylation site"
 FT Modified-site 200..202 /note="Protein kinase C phosphorylation site"
 FT Modified-site 201..204 /note="CAMP and cGMP-dependent protein kinase phosphorylation site"
 FT Modified-site 215..218 /note="Casein kinase II phosphorylation site"
 FT Modified-site 219..222 /note="Casein kinase II phosphorylation site"
 FT Modified-site 225..228 /note="Casein kinase II phosphorylation site"
 FT Modified-site 230..233 /note="Casein kinase II phosphorylation site"
 FT Modified-site 275..278 /note="N-glycosylation site"
 FT Modified-site 279..282 /note="N-glycosylation site"
 FT Modified-site 307..312 /note="N-myristoylation site"
 FT Modified-site 329..335 /note="Tyrosine kinase phosphorylation site"
 FT Domain 383..540 /note="Sterol-sensing domain, SSD, domain"
 FT Modified-site 418..422 /note="N-myristoylation site"
 FT Modified-site 504..509 /note="N-myristoylation site"
 FT Modified-site 508..510 /note="Protein kinase C phosphorylation site"
 FT Modified-site 530..533 /note="N-glycosylation site"
 FT Modified-site 535..540 /note="N-myristoylation site"
 FT Modified-site 541..544 /note="Amidation site"
 FT Modified-site 561..563 /note="Protein kinase C phosphorylation site"
 FT Modified-site 572..575 /note="Casein kinase II phosphorylation site"
 FT Modified-site 597..600 /note="Casein kinase II phosphorylation site"
 FT Modified-site 662..664 /note="Protein kinase C phosphorylation site"
 FT Modified-site 678..681 /note="N-glycosylation site"
 FT Modified-site 681..688 /note="Tyrosine kinase phosphorylation site"
 FT Modified-site 692..695 /note="N-glycosylation site"
 FT Modified-site 737..740 /note="N-glycosylation site"
 FT Modified-site 740..743 /note="Casein kinase II phosphorylation site"
 FT Modified-site 746..748 /note="Protein kinase C phosphorylation site"
 FT Modified-site 759..761 /note="Protein kinase C phosphorylation site"
 XX EP1229046-A2.
 XX 07-AUG-2002.
 XX 28-JAN-2002; 2002EP-00001167.
 XX 30-JAN-2001; 2001WO-US000663.
 PR 30-JAN-2001; 2001WO-US000664.
 PR 30-JAN-2001; 2001WO-US000665.
 PR 30-JAN-2001; 2001WO-US000667.
 PR 30-JAN-2001; 2001WO-US000668.

PR 30-JAN-2001; 2001WO-US000669.
 PR 23-MAY-2001; 2001US-00864761.
 PR 09-OCT-2001; 2001US-0327898P.
 PA (AEOM-) AEOMICA INC.
 XX Zhan J;
 PT WPI; 2002-676582/73.
 DR N-PSDB; ABV78761, ABV78762.
 XX Novel isolated human testis expressed Patched like protein (HTPL), useful for identifying agonist and antagonist and specific binding partners, and for treating subjects having defects in HTPL.
 XX Claim 13; Fig 4; 718pp; English.
 CC The present sequence is the protein sequence for human testis expressed Patched like protein (HTPL) short form (HTPL-S). HTPL has two isoforms, with a few single base pair differences between the two. One of the CC single base pair changes introduces a premature stop codon in HTPL-S (S for short) compared to HTPL-L (L for long). HTPL shares an overall CC structure organisation with the Patched protein. The shared structural CC features strongly imply that HTPL plays a role similar to that of CC Patched, and is a potential tumour suppressor. HTPL is important in CC regulating male germ cell development, and the HTPL gene was mapped to CC human chromosome 10p12.1. HTPL and its coding sequence are useful for CC diagnosing a disorder caused by mutation in HTPL, and in therapy and CC manufacture of a medicament for treatment or prevention of such disorder CC associated with decreased expression or activity of human HTPL. Such CC disorders include disorders of testis, or adrenal, adult and foetal CC liver, bone marrow, brain, kidney, lung, placenta, prostate, skeletal CC muscle or colon function. HTPL proteins and nucleic acids are clinically CC useful diagnostic markers and potential therapeutic agents for male CC infertility and cancer
 XX Sequence 767 AA;
 SQ
 Query Match 5.1%; Score 351.5; DB 5; Length 767;
 Best Local Similarity 20.5%; Pred. No. 1.1e-22;
 Matches 158; Conservative 137; Mismatches 285; Indels 189; Gaps 31;
 QY 334 HT-----LIGQFFQGGTGWASWPLTILVLSVIPVVALAAGLVF--TELTTPDELWSAP 386
 DB 115 HTDCLGLLSRTFQWLGWQVGAHPWIFLLAPLMLTAALGTGFLYLPKDEEDLEHYTPV 174
 QY 387 NSQARSEKAFHDQHFQPPFRNTQVILTAPNRSS-----YRDSLLLPKPNPSGIL 436
 DB 175 GSPAKAERRFVQGH---FTTNDYRFSASRRSTEANFVSLVSVSYSDSLDLPATPAEVS 230
 QY 437 DLDLLELLELQERLRLHLQVMSPEAQRNLSLODIC--YAPLNPNTSLYDCCI--NSLLQ 492
 DB 231 KLDGAVQDLRV-----AREKSGQIQYQVCARYAL-----CVPPNPLY 270
 QY 493 YFQNNRTLLLTANOTLMGQTSQVDKQHFYCANAPLTFKDGTTALALS CMADYGA-PVF 551
 DB 271 AMQVKNLTNL-----SSISFPAY-----NHGRHPLY 296
 QY 552 PFLATGGY-----KGKDYSEABALIMTFSLNNTPAGDPRL-AQAKLWEAFLEMRAP 603
 DB 297 LTGPFPGYILGSLGMGQLLRAKAMRLLYLK---TEDPEYDVQSKQWLTHLLDQFTNI 353
 QY 604 QRRMA--GMFQVTFTA-ERSLEDEINRTTAEDLPATSVYIVFLYISALSGSYSSMSRV 660
 DB 354 KNILAKKIEVHFHTSLRQLEFATSVTV--IPVFLAYILILIFAVTSCFRFD----- 406
 QY 661 MYDSKATLGLGVAVVLGAVMAAMGPFPSYLGSRSSILVILQVVPFLVSLVGADNIFIFVLE 720
 DB 407 CIRNKMCAVAFGVISAFLLAVSGFGLLHIGV-PFVILVANSFPLILGVGDVDMFIMISA 465
 QY 721 YORLPRRPGBPBEVHIGRALGRVAFSMLCSUSEAICPFLGALTMPAPVTRTALTSGLAV 780
 DB 466 WHK-TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSQVCFIYTGMTL 523

```
CC 781 ILDFLOMSAFVALISLDSKQEARSLDVCCVKPQELP-----PFGQG 824
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 524 LFCYFNTCTCGAFWALDGKEV-----VCLWLKKAADPKWPSFKKCPFFGVPDEHG 578
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 825 EGL--LLGFFQKAYAPFLHMTTRGVLLLFALFGVLSYMSCHISVGLDQELALPKDSY 882
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 579 TDIHPSLFFRDYFGFPLTRSSKVFVFIYLYLISSYIGCFHVQEGLDLRNLASDDSY 638
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 893 LLDYFLFLNRYE-EVGAPYFVTTLGYNFSSAGNAICSSAGCNPFSTQKIOVATFP 941
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 639 IPTFVENYENYSDYGPVMVITVKVDWIK-----DVRQKLENTKIF 683
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 942 EQSYLAIP--ASSWDDFIDMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMG 999
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 684 EKNVVDKNLTEFWLDAYVQ-----YLGNSQD---PNEKNT-----717
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 1000 SVRPSVEQHKYLPFLNDRPNIKCPKGLAAYSISVNLTSQGVLASR 1048
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
CC 718 -----PMNNIPDFLSNFPN-----FQHDINISSNEIISR 748
CC      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

RESULT 98

```
AD121155
ID AD121155 standard; protein; 783 AA.
XX
AC AD121155;
XX
DT 15-APR-2004 (first entry)
XX
DE Novel human protein #130.
XX
KW forensic; nutritional source; damaged tissue; diseased tissue;
KW myeloid cell disorder; lymphoid cell disorder;
KW bone cartilage tissue growth; tendon tissue growth;
KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;
KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.
XX
OS Homo sapiens.
XX
PN WQ2003025148-A2.
XX
PD 27-MAR-2003.
XX
PF 19-SEP-2002; 2002WO-US029964.
XX
PR 19-SEP-2001; 2001US-0323739P.
XX
PR 13-SEP-2002; 2002US-00323739.
XX
PA (HYSE-) HYSEQ INC.
```

```
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
XX Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
XX Haley-Vicente D;
XX
XX WPI; 2003-354603/33.
XX N-PSDB; ADI21871.
XX
XX New polynucleotides and secreted proteins, useful for treating myeloid or
XX lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
XX tissue growth or regeneration, in wound healing, and in tissue repair and
XX replacement.
```

```
XX Claim 20; SEQ ID NO 406; 156pp; English.
XX
XX The invention relates to an isolated polynucleotide encoding a
XX polypeptide with biological activity. The polynucleotides and
XX polypeptides are useful in diagnostics, forensics, gene mapping,
XX identification of mutations responsible for genetic disorders and other
XX traits, to assess biodiversity, as nutritional sources or supplements.
XX The polynucleotides may also be used as molecular weight markers,
XX chromosome markers or map related gene positions, or as an antigen to
XX raise anti-DNA antibodies or elicit immune response. The polypeptides are
```

```
CC useful for raising antibodies, as markers for tissues in which the
CC corresponding polypeptide is expressed, for re-engineering damaged or
CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
CC bone cartilage, tendon, ligament and/or nerve tissue growth or
CC regeneration, in wound healing, in tissue repair and replacement, in
CC healing of burns, incisions and ulcers, and in treating cancer. The
CC present sequence represents the amino acid sequence of a novel human
CC protein.
```

SQ Sequence 783 AA;

Query Match 5.1%; Score 349; DB 7; Length 783;

Best Local Similarity 21.5%; Pred. No. 1.9e-22;

Matches 193; Conservative 160; Mismatches 335; Indels 208; Gaps 42;

QY 444 LLEQLERLHLQVWSPQAQNI SLQDICYAPLNPDTSLYDCINSLQYFQN-----NRT 499

Db :

Db 10 ILKLLAAVTKIQV--PRGFNYTPAHIC--LNNDKTCIADDIVH-VLEELKNAARNTRT 64

QY 500 LLLLTANQTLMGQTSQVDKDHFLYCANAPLT-PKDGTAALALSCMADYGAVFPFLAIGG 558

Db :

Db 65 NFAIT-----YFTHLKDGRA-----VYNGHQLGG 89

QY 559 --YKGD--YSEAEALIMTFLNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGFQVTF 615

Db :

Db 90 VTVHSKORVKSAAEQITLYTLQSLNSLNDMVAER--WESSFCDTVRLFQKSNKYOMYPY 147

QY 616 TAERSLEDEINRTAEDLPFPATSYIVIFLYISALGYSWSRVWVDSKATLGLGGVAV 675

Db :

Db 148 TSS-SLREDFQKTSRVSERYLVTSLLV---VTMAILCCS--MODCKRSKWLGLGLVLT 201

QY 676 VLGAVMAMGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIVLEYQLRPRPGPREVH 735

Db :

Db 202 ISLATLTHAGIINLTGGKYNSTFLG-VPPVMLGHLYCTFEMLSW-----RKTREDQHKV 256

QY 736 IGRALGRVAPSKMLCSSEA---ICFFLGA--LTPMPARVTFALTSGLAVILDFLLQMSA 790

Db :

Db 257 -ERTAADVADSMLSFSLTAMVLTFGIGASPFNIEAARIFCCNCSIAIFPNLYLVLSF 315

QY 791 FVALLSLDSKQEARSLDVCCVKP-----QELP-----PFGQE-----825

Db :

Db 316 YGSSLVFTGYIENNYQHSIFCRKVPKPEALQEKPAWYRFLITARFSETAEGETEYES 375

QY 826 GLLLGFFQKAYAPFLHMTTRGVLLLFALFGVLSYMSCHISVGLDQELALPKDSYLLD 885

Db :

Db 376 HLLVCFKRYCDWITNTYVVPFVLYIYISFALMGVQLVQVSEGSLSNIVATATQITIE 435

QY 886 YFLFLNRYFVGAPYFVTTLGYNFSSAGNAICSSAGCNPFSTQKIOVATEPEQSY 945

Db :

Db 436 YTTAQQKYFSNYSPIV-----GFYIYESIEYNTSVQEDV 470

QY 946 LAIPAS-----SWDDFIDMLTPSSCCRLYIS-GPNKDKFCPSTVNSLNCNCSITMGS 1000

Db :

Db 471 LEYTKGFVRISWFESYLYNLR-----KLVSTGLPKKFTDMLNSF--LK-----514

QY 1001 VRPSVEQHKYLPFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRMAHKPKL-NS 1059

Db :

Db 515 -----APQFSHQEDIIFSK-----KYNDEVD-----VVASRMFLVAKTMEYTR 553

QY 1060 QYTBALRAARELANITADLRKVGTPDPAPEVFPYITNNVFEQVYLTLPGLFMLSJC 1119

Db :

Db 554 BELYDLLETLRRL--SVTSKVPFIV-FNPSF-----VYMDRYASSL--GAPLHNSC 599

QY 1120 LVPTFAV--SCLLGLDLRSGLNLN---LSIVMLVDTVGFMALMDISNVAISLINLVA 1174

Db :

Db 600 ISALFLFFSAFLVA-----DSLINWITLVTVSVEGVIGFWLTKWVELDCISVLCLYIG 655

QY 1175 VGMSEFVSHITRSFAIS---TKPTWLEPRAKATISMGSAVAGVAMTN-----LPGILV 1226

Db :

Db 656 INVTIDNCAPMLSTFVLGKDFTRTKWKNALV-----HGVAILQSYLCYIVGLPI 706

QY 1227 LGLAKAQLQIQTFFRLNMLITLLGLLHGLVFLPVILSVYVGPVNPALAEQKRAE 1282

Db :

Db 707 LAAPVSNL-TCTLFRCLFLIAFVTFHCFAILPVILTFPLPPS-----KKRCKEK 754
RESULT 99
ID ADA54204
XX ADA54204 standard; protein; 783 AA.
AC ADA54204;
XX 20-NOV-2003 (first entry)
DT Human protein, SEQ ID 1772.
DE
KW Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.
XX Homo sapiens.
XX
XX EP1293569-A2.
XX 19-MAR-2003.
XX 21-MAR-2002; 2002EP-00006586.
XX 14-SEP-2001; 2001JP-00328381.
XX 24-JAN-2002; 2002US-0350435P.
XX (HELI-) HELIX RES INST.
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PI Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahara K, Masuho Y;
XX WPI; 2003-395539/38.
DR N-PSDB; ADA52565.
XX
XX New polynucleotides encoding full-length polypeptides, e.g. secretory
XX and/or membrane proteins, useful for developing medicines for diseases in
XX which the gene is involved, or as target molecules for gene therapy.
PS Claim 14; SEQ ID NO 1772; 205pp; English.
XX
XX The present invention relates to novel human secretory or membrane
XX proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
XX ADA54071). The coding sequences are useful in the gene therapy of
XX diseases caused by abnormalities of the proteins, e.g. cancer,
XX inflammatory diseases, osteoporosis or neurological disease.
XX
XX Sequence 783 AA;
Query Match 4.9%; Score 339; DB 6; Length 783;
Best Local Similarity 21.1%; Pred. No. 1.6e-21;
Matches 190; Conservative 161; Mismatches 333; Indels 216; Gaps 41;
QY 444 LLEQLERLHLQVWSPQARNISLQDICYAPLNPNTSLYDCCINSLLQYFON----NRT 499
Db 10 ILKHLAAVTKIQV--PRPGNTFFAHICI--LNNDKTCIVDDIVH-VLEELKNARATNR 64
QY 500 LLLLTANQTLMGQTSQVQWKDHFYCANAPLT-FKDGTAALALSCMADYGAPVFPFLAIGG 558
Db 65 NPAIT-----YPITHLXKGRA-----VYNGHQLGG 89
QY 559 --YKGD-YSEBALIMTSLNNYPAGDPRLAQAKLWEAPLEEMRAFORRMAGPQVTF 615
Db 90 VTVHSDRVKSAEAIQLTYLQINSINDMVAER--WESSFCDTVRLFKQSNKVMKVPY 147
QY 616 TAERSLEDEINTTAEDLPFATSYIVILYSLALGSYSSRWVDSKATLGGGVAV 675
Db 148 TSS-SLRDFQKTSRVSRYLVTSLIV---VTMAILCCS--MQDCVRSKPWLLGLVLT 201
QY 676 VLGAVMAMGFFSYLGRSSLVILQVVPFLVSLVGADNIFIVLBYQRLPRPGPREVH 735

Db 202 ISLATLTAAIGIINTGGKYNSTFLG-VPFVVLGHGLYGTFFEMLSW-----RKTREDQHVK 256
QY 736 IGRALGRVAPSNMLLCSLEA---ICPFELGA--LTPMPAVRTEALTSGLAVIDLFLQMSA 790
Db 257 -ERTAAVYADSMLSFSLTTAMVLYTFGIGASPTNIEARIFCCNSCAIFNFVLVLSF 315
QY 791 FVALSLDSKRQEASRLDVCCCKP-----QSLP-----PFGQGE-----825
Db 316 YGSSLVFTGYIENNYQHSIFCKRKVPKPEALQEKPAWYRFLLTARFSEDTAECEANTYES 375
QY 826 GLLGPFQKAYAPFLHWTITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLD 885
Db 376 HLLVCLKRYCYCDMTNTYVFPVFLYLIYSPALMGVLYQVSEGSLSNIVATATQTIE 435
QY 886 YFLPLNRYVEGAPV---YFVTLGYNFSS-----EAGMNAICSSAGCNCNFSFTOKIQY 936
Db 436 YTTAQQKYPNSYSPVIGFYIYESIYNTSVQEDVLEYTKGFRISRFSYLYNLKLV 495
QY 937 ATEPPEQSYLAIPASSWVDDFDIMLTPSSCCRLYISGPNKDKFCPSTVNSLNLKNCMSI 996
Db 496 STGLPKK-----NPTDMLRNS-----FLKAPQFSHFQSDIIFS-----528
QY 997 TWGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAYSTVNLTSDGVLASRPMAYHKPL 1056
Db 529 -----KKY-----ND-----EYDVVASRMFLVAKTM 549
QY 1057 K-NSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYTTITNPFYEQYLILPEGLFM 1115
Db 550 ETNREELYDLLETLRRL--SVTSKVKFIV-FNPSP-----VIMOKIASL--GAPL 595
QY 1116 LSLCLVPTFAV--SCLLGLDLRLSGLNL---LSIVMILVDTVGFMALWDISYNAVSLIN 1170
Db 596 HNSCISALFLFPFSAFLVA-----DSLINVWITLVSVSEFVGIVGFWTLKKVELDCISVLC 651
QY 1171 LVSAGVMSVEFVSHITRSPAIS---TKPTWLERAKEATISMGSAVPAVAGVMTN----LP 1222
Db 652 LIYGINYTIDNCAPMLSAFLGKDFTRTKWKNALV-----HGVAIIQSYLCYIV 702
QY 1223 GILVLGLAKAQLIQIFPERNLILTLGLLHCLVPLVTLVYGVDPVNPALALEOKRAEE 1282
Db 703 GLIPLAAPSNL-TCTLFRCLFLIAFVTFHCFAILPVILTFPLPPS-----KKRCKEK 754
RESULT 100
ADD46676
ID ADD46676 standard; protein; 608 AA.
XX
AC ADD46676;
XX
XX 02-DEC-2004 (revised)
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein AAC99398, SEQ ID NO 12361.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
OS Unidentified.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:41:11 ; Search time 53 Seconds
(without alignments)
2418.127 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAIISNPLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

PIR 80:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2385	34.5	1278	2 T30188	Niemann-Pick C dis
2	1394	20.2	1055	2 T05663	hypothetical prote
3	1347	19.5	1456	2 T15961	hypothetical prote
4	1329	19.2	1170	2 S52525	probable membrane
5	646.5	9.4	1434	2 T30172	transmembrane prot
6	628.5	9.1	1442	2 T18538	patched protein -
7	600	8.7	915	2 S44797	F09G8.4 protein -
8	598.5	8.7	933	2 T25600	hypothetical prote
9	583.5	8.4	1182	2 T13952	membrane protein p
10	573.5	8.3	1003	2 T26746	hypothetical prote
11	572	8.3	1220	2 T18291	patched protein -
12	539.5	7.8	889	2 T29590	hypothetical prote
13	514	7.4	1405	2 T15830	hypothetical prote
14	473.5	6.9	1015	2 T15830	hypothetical prote
15	464.5	6.7	1299	1 S08119	membrane protein p
16	458.5	6.6	955	2 T21612	hypothetical prote
17	436	6.3	956	2 A89153	hypothetical prote
18	409.5	5.9	936	2 T26521	hypothetical prote
19	407	5.9	800	2 T26683	hypothetical prote
20	399.5	5.8	881	2 T31739	hypothetical prote
21	381	5.5	840	2 T33217	hypothetical prote
22	379	5.5	714	2 T16126	hypothetical prote
23	367	5.3	690	2 T23399	hypothetical prote
24	367	5.3	877	2 T24097	hypothetical prote
25	349	5.1	845	2 T25657	hypothetical prote
26	347.5	5.0	820	2 T32908	hypothetical prote
27	312	4.5	890	2 T22186	hypothetical prote
28	297.5	4.3	871	2 T28706	hypothetical prote
29	294	4.3	983	2 T21213	hypothetical prote

30	276.5	4.0	633	2 S44795	F09G8.3 protein -
31	259.5	3.8	413	2 S28276	hypothetical prote
32	224	3.2	1276	2 T18526	SREBP cleavage act
33	189.5	2.7	1154	2 T48829	related to SREBP c
34	187.5	2.7	932	2 T28820	hypothetical prote
35	186.5	2.7	1227	2 T20370	hypothetical prote
36	179.5	2.6	787	2 H71453	hypothetical prote
37	173	2.5	969	2 T33156	hypothetical prote
38	170	2.5	746	2 A75018	transport protein
39	154	2.2	823	2 B81282	probable integral
40	149	2.2	1011	2 T07712	probable ABC-type
41	149	2.2	1051	2 AG3455	multidrug resistat
42	148.5	2.1	746	2 H84301	hypothetical prote
43	146	2.1	1041	2 AC0423	multidrug efflux p
44	143	2.1	724	2 H69780	antibiotic transpo
45	141.5	2.0	1049	2 AP0561	acriflavin resista
46	140.5	2.0	1027	2 D85757	probable efflux pu
47	140.5	2.0	1027	2 H90861	probable multidrug
48	140.5	2.0	1046	2 S39630	multidrug-efflux t
49	140.5	2.0	1046	2 E83593	RND multidrug effl
50	139	2.0	1053	2 D83206	probable RND efflu
51	138	2.0	1057	2 AH2198	RND multidrug effl
52	136.5	2.0	1032	2 T20342	hypothetical prote
53	135	2.0	1067	2 G81825	probable drug effl
54	135	2.0	1067	2 E81051	multiple transfera
55	134	1.9	749	2 H91170	hypothetical membr
56	133	1.9	750	1 D69403	conserved hypotet
57	132	1.9	736	1 C69307	conserved hypotet
58	131	1.9	1166	2 C96945	cation efflux syst
59	130.5	1.9	808	1 VGBESM	glycoprotein B pre
60	130.5	1.9	1036	2 D82357	probable multidrug
61	130	1.9	772	2 H86016	hypothetical prote
62	129	1.9	801	2 A98862	Na+/H+ antiporter
63	129	1.9	1043	2 AB0370	multidrug efflux p
64	127.5	1.8	1045	2 D83393	RND multidrug effl
65	127	1.8	1036	2 F82263	probable multidrug
66	127	1.8	1061	2 S75508	acriflavin resista
67	127	1.8	1093	2 H84126	cation efflux syst
68	126.5	1.8	1013	2 G87236	conserved integral
69	126	1.8	674	2 D72329	hypothetical prote
70	126	1.8	1047	2 G87398	AcRb/AcrD/AcrF fam
71	125.5	1.8	749	2 A75560	conserved hypotet
72	125.5	1.8	1172	2 D72482	probable antibioti
73	124	1.8	1025	2 A83186	probable RND efflu
74	123.5	1.8	793	2 C83260	hypothetical prote
75	123.5	1.8	964	2 E70826	probable membrane
76	122.5	1.8	676	2 G69154	conserved hypotet
77	122	1.8	826	1 Q0BE11	ribonucleoside-dip
78	121.5	1.8	385	2 T52483	hypothetical prote
79	121.5	1.8	477	2 H91123	probable oxidoredu
80	121.5	1.8	477	2 G85968	probable oxidoredu
81	121.5	1.8	1024	2 A83557	acriflavin resista
82	121.5	1.8	1037	2 A80816	probable efflux pu
83	121.5	1.8	1043	2 H83071	RND multidrug effl
84	120.5	1.7	694	2 E72272	conserved hypotet
85	120	1.7	2149	2 C96695	ribulose biphosph
86	119.5	1.7	775	2 A83386	protein translocas
87	118.5	1.7	1075	2 S76433	cation efflux syst
88	118.5	1.7	705	2 T35165	probable integral
89	118.5	1.7	753	2 T32844	hypothetical prote
90	118.5	1.7	1029	2 D83120	probable RND efflu
91	118.5	1.7	1036	2 B83466	probable RND efflu
92	118	1.7	368	2 C71131	hypothetical prote
93	117.5	1.7	767	2 D70131	conserved hypotet
94	117.5	1.7	1049	2 G85543	acridine efflux pu
95	117.5	1.7	1049	2 C90693	acridine efflux pu
96	117.5	1.7	1049	2 B36938	acriflavin resista
97	117.5	1.7	1485	2 S23756	CPR protein - Afr
98	117	1.7	631	2 B98137	hypothetical 46.1K
99	117	1.7	1969	2 T38495	hypothetical prote
100	116	1.7	1050	2 A50380	multidrug efflux p

ALIGNMENTS

RESULT 1

T30188
Niemann-Pick C disease protein - mouse
N:Alternate names: NPC1 protein
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
R:Accession: T30188
R:Loftus, S.K.; Morris, J.A.; Carstea, E.D.; Gu, J.Z.; Cummings, C.; Brown, A.; Ellison, Science 277, 232-235, 1997
A:Title: Murine model of niemann-pick C disease: mutation in a cholesterol homeostasis g
A:Reference number: Z20765; PMID:97362324; PMID:9211850
A:Accession: T30188
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-1278 <LOF>
A:Cross-references: UNIPROT:O35604; UNIPARC:UPI00000299BC; EMBL:AF003348; NID:g2251241;
C:Genetics:
A:Gene: Npc1
A:Map position: 18

Query Match 34.5%; Score 2385; DB 2; Length 1278;
Best Local Similarity 40.2%; Pred. No. 1.2e-158;
Matches 525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;
QY 14 LLLRLAQSEPYTHIQGYCAFYDCGKGNPELGSMLTNSVCLSNTPARKITGDHIL 73
DB 12 LLLLLCPAQVFSG-----SCWYBGC-----IATGD-----KRYNCKYSGPPKPLPKDGYDL 59
QY 74 LQKICPRLTYGTQACSAKOLVSEASLSITKALLTRCPACSNFVNLHCHNTCSNQ 133
DB 60 VQELCPGLFF-DNVSLLCCDIOQLTKSLQLPLQFLSRCPSCFYNLMTLFCELTCSHQ 118
QY 134 SLFINVTRVAQLGAGQLPA---VWAYEAFYQHSFAEQSYDCSRVRVPAAATLAVGTWC 189
DB 119 SQFLNVTATEDYFDKTPENKTNVKELEYVYVQGSFANAMYNACRDEAPSSNEKALGLLC 178
QY 190 GYVGSALCNAQRLNFGQDTGNGLAPLDI-----TFHLEPGQAVGSGIQPLNEGVARN 244
DB 179 GRDARA-CNATNWIETMFKNDNGQAPFTIIPVSDLSIL-----GMEPRNATKGCN 229
QY 245 ESQDDVATCSCDCAASC-----PAIARPAQLDSTF-----VLGQMPGSIV 286
DB 230 ESDVETGTCSCDCCSIVCGPKQPPPPMPWRINGLDAMVIMVYVAFILVFVFGALL 289
QY 287 LIILCSFPAVVTILLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQFFQGWGT 346
DB 290 AVWCHRRYFVSEYTPIDSNIAFVNSS---DKGEASCCDPLGAADFDDCLRRMFTKGA 345
QY 347 WVASWPLTILVLVSVIPVVALAGLVTELTDPVELWSAPNGQARSEKAFPHQHGPPFR 406
DB 346 FCVRNPTCIIFFSLAFITVCSSGLGVFQVTTNPVELWSAPHQARLEKEYFDKHGPPFR 405
QY 407 TQNVILTAPNRSRYVDSLLGLPK-NFSGILDLALLELELQERLHRLQVWSPEAORN- 464
DB 406 TEQLITQANTSVHIIYEPTPAGADVFGPPLNKEILHQLVNLQ-----TAISITASYNN 460
QY 465 --ISLQDQICVAPLNPNSTSLDYCCINSLLQYFQNNRTLLLTANTQTLMGQTSQV----- 516
DB 461 ETVTLQDQICVAPLSPYNK---NCTINVLNYPQNSHAVL-----DSQVGDGDFYI 506
QY 517 --DWKDHLYCANAPLTFKDGATALASCHADYCAPVFPFLAIGGYKGYKDYSENAALIMTF 574
DB 507 YADYTHFLYCVRAPASLNDTSLHGCPLGTGCGPVFPMLVGLGDDQYNNATALVITF 566
QY 575 SLANNYPAGDRLAQAKLWEAFLEMRAPFORMMAGMFQVTFTAERSLEDEINRTTAEDLP 634
DB 567 PVNNYNDTERLQANAWKEKIFSFVKYKN---PNLTISFTAESIEDELNESDVF 623
QY 635 IPATSYIVIFLYLSALGSYSSRVMVDSKATLGLGGVAVILGAVMAAGFFSYIGIRS 694

DB 624 TVIISYVMEFLYISLALGHITQCSRLIVDSKISLIGIAGILIVLSSVACSLGIFSYMGMPL 683
QY 695 SLVILQVPELVLSVGADNIFIVLEYOQLPRRPGEPREVIHGALGRVAPSMCLLSLSE 754
DB 684 TLIVIEVIFPLVAVGVDNIFILVQTYQDRERLQEBTLDQOGLKEILGEVAPTMELSFSE 743
QY 755 AICFFLGAITPMPAVRTFALTSLGLAVILDFELLOQMSAFVALLSDSKRQEARLSDVCCVK 814
DB 744 TSAFFFGALSMSPAVHTFSLPAGNAVILDFLLQITCFVSLGLLDIKQEKHNLIDLCVR 803
QY 815 PQELPPPGQ-----EGLLGFFQKAYAPFLHMITRGVLLFLALFGVLSYMSCHISVG 870
DB 804 GAD---DQCGSHASESYLFFPKYFAPLLKDWLRPIVAVFVGLVSFSAVVNVKVDIG 860
QY 871 LDQELALPKDSYLLDYFLNRYFEVCAVYFVTTGLYNFSSSEAGMAICSSAGCNPFSF 930
DB 861 LDQSLMPNDSYVIANFKSLAQYLHSGPPVTFVLEEGYNSRRKQGNMVCGMGCDNDL 920
QY 931 TQKIQVATEPPEQSYLAIPASSWDDFIDMLTP-SSCCRLYISGPNKDKPCPSTVNSLNC 989
DB 921 VQQIFNAELDTYTRVGPAPSSWIDDYFDMVSPQSSCCRLY---NVTHQFCNASVMDPTC 977
QY 990 LKNCMSIT-MGSVRPSVEQFHKYLWFLNDRPNIKCPKGLAAYSTSNLTSDDQVL-AS 1047
DB 978 VR-CRPLTPEGKQRPQCKEPMKFLPMLFSLDNPNPKCKGKGHAAYGSAVINVGDDTYIGAT 1036
QY 1048 RFMAYHKPKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYTIITNVFYEQYL 1107
DB 1037 YFTYHTILKTSADYTDAMKARLIASNIETWMS-KGSD--YRVFPYSVFVFEQYL 1093
QY 1108 ILPEGLFMLSCLVPTTFAVSCLLGLDLRSLGLNLLSIVMLVDTVPGMALWDISYNAV 1167
DB 1094 IIDDTIFNLSVLSIGSIFLTVLVGLCELSAVIMCITAMILVNMFGVMWLWGLSLNAVS 1153
QY 1168 LINLVSAGMSVEFVSHITRSFALSTKPTWLERAKEATISGSAVAGVAMTNLPGIIVL 1227
DB 1154 LVNLVMSCGISVEFCSHITRAFTMTKGSRSVRAEALAHMGSSVFSGITITKFGGIVVL 1213
QY 1228 GLAKAQLQIIFFRNLNLLITLLGLLGLVFLPVLISYVGPVWPA 1272
DB 1214 APAKSIQEIFFRMYLAWLLGATHGLIFLPLVLSYIGPSVNA 1258

RESULT 2

T05663

hypothetical protein F22113.120 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004

C:Accession: T05663

R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15420

A:Accession: T05663

A:Molecule type: DNA

A:Residues: 1-1055 <BEV>

A:Cross-references: UNIPROT:Q9SVF0; UNIPARC:UPI00000A52EC; EMBL:AL035539

A:Experimental source: cultivar Columbia; BAC clone F22113

A:Genetics:

A:Map position: 4

A:introns: 24/3; 60/3; 99/3; 150/1; 193/1; 216/3; 278/3; 297/2; 336/3; 364/3; 396/3; 411/3

A:Note: F22113.120

Query Match 20.2%; Score 1394; DB 2; Length 1055;

Best Local Similarity 28.5%; Pred. No. 2.6e-89;

Matches 377; Conservative 221; Mismatches 385; Indels 340; Gaps 41;

QY 36 YDECGNPESLSMTLSNVCSLSTNPARKITGDHLI--LLQKICPRLTYGTNTQACCSA 93

DB 2 YDICHRSR--GKVL-----NCPYASPS--IQDELFSAKIQSLCPTI-----SGNVCCTE 48

QY 94 KQLVSLSEASLSITKALLTRCPACSNFVNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAV 153

Db 49 TQPDTLRSQVQAAPFLVCPACLRNPLNLFCELSQSPNQSILFINVTSVAEV-SGNL-TV 106
Qy 154 VAYEAFYQHSFAEQSDSCSRVRVPAATAATLAVGTMGVYGSALCAQRLNFGQD---TG 210
Db 107 DGDVHITDTFGEGLYESCKEKFQGMTRAINFV---GGGAKNFREWFTFIGQKAPSG 162
Qy 211 NGLAPLDITFHLLEPQAVSGIOPLEGVARGNESQGDVATCSCQDCAACPAIARQ 270
Db 163 FPGSPYAINFKSIP---BSSAMVPMN-----VSYSACACSPPEPLPH 203
Qy 271 ALDS-TFYLG-----OMPGSLVLIILCSVFAVVTTILLGFRVAPARDKSMVDPKKG 322
Db 204 DEDSCSIRIGPLKVCRIELSMALVYVLLVSCFEGWAGLNRRNTTQPLDSSKPLHPVEE 263
Qy 323 TSLSDKLSFSTHTLLG-----QPFQGMGTWVASWPLTILVLSVIPVA 365
Db 264 DGINSEMK---ENILGVKQVORHAQLSPQVQRYMAKPYRSYGSMIARNPSLVLFMSVAIVLA 320
Qy 366 LAAGLVFTLTDPVELWAPNSQARSEKAFHDQHPGFPFRTNQVIL-TAPNRSSVRYDS 424
Db 321 LCSGLYNFKVETREPKLWGPESKAABEKKFFDTHLSPFYRIEQLTATVPDPKSGRAPS 380
Qy 425 LLLGPNFSGILDLDLLLELLELOERLRLHQLWMSPEAQRNISLQDICYAPLNPDTNSLYD 484
Db 381 I-----VIDENILL-LFDIQK----- 396
Qy 485 CCNISLLOYFQNNRTLLLTANTQMTMGQTSQVDWKDHFYCANAPLTFKDGFTALALSCMA 544
Db 397 -----YFKMD-----SGTDDYGGVEHAEC-----FOHYTS-SETCLS 429
Qy 545 DYGAVPFPPLATGVYKGDYSE-----AELIMTFSLNVP 580
Db 430 AFQAPVDPASVAVLGGFSGNNYSVMSELGCVSPFDCYSVDRKRTLFOATAPVVTYPVNNV- 488
Qy 581 AGDP--RLAOKLWEAEAL---EEMRAFORWAGMFOVTTAERSLEDEINRTTAEDLP 634
Db 489 IGDSSNENARAVAKESFQLAKBELLPMV--SKNLSFSFSSSIEBELKRESTADVI 546
Qy 635 IFATSYVIFLYISALGSSWSRVMVDSKATLGLGVAVVLGAVMAAMGPFSLGTRS 694
Db 547 TTAASVLMFVIVSVTLGADAPQFYFYSKVLGLSGVVLVLLSVLGSVGVFSALGVKS 606
Qy 695 SLVILQVFPFLVSGADNIFIVLEYQRLPRRPGEPREHVHIGRALGRVAPSMMLCSLSE 754
Db 607 TLIIMEVTPFLVAGVDMNCILHVAKQPREVS--LEQRISALVEVGPSTLASLSE 664
Qy 755 AICFFLGALTMPAVRTFALTSLGLAVILDPILOMAFVALLSLDSKROEASRLDCCCVK 814
Db 665 VLAFAVGAEPMPACRIFSMFAALAIMLDFLQITAFVALIVDFCKRSADNRIDCFPCIK 724
Qy 815 ----POELPPGOGGLIGPFQKAYAPFLLHWITRGVVLLLFALFGVSLYSMCHISVG 870
Db 725 VPSSRESVEGREGPFLERYKMEVHAPVLGLWGVWVAVVFA----- 769
Qy 871 LDOELALPKDSYLLDYFLFLNRYFEVGAPEVYFVTTLGYNFSSAGMNAICSSAGCNFSP 930
Db 770 -----PAL 772
Qy 931 TOKIQATEFPQSYLAIPASSWDDFDLWLPSS---CCRLYISGNKDKKPCSTVNSLN 988
Db 773 ASIIIRASQASITSYIAKPAASWLDLDFLWLSPEAFGCCCKFTNG-----SYCPP---DDQ 825
Qy 989 CLKNCMSITMGSVRPSVEOFKHLFWLNDRNKICPKGLAAYSTSVNLT--SDGOVLA 1046
Db 826 CFRHS---DLVQDRSTAGFRKLPWFLNALFSADCAKGHGYNSVDLKGYESGVIOA 882
Qy 1047 SRFMAYHKPKNSQDYTEALRAARELANITADLRKVPDTPAFEVPPYTIITNVFEQYL 1106
Db 883 SEFRTVHTFL-NTQ-----IDIFFYSVYVIFFEQYL 912
Qy 1107 TILPGLFMLSCLVPTFAVSCILLGLDLSGLLNLISVIMILVDVGMALWDISYNV 1166
Db 913 NIWTVALTNLATAIV-----GIQLNAV 934

Qy 1167 SLINLVSAVMSVRFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLFGILV 1226
Db 935 SVVNLIMSIGIAVEFCVHISHAFVLMSSGDR-BHREAREALETMGASVFSGTLTKLVGVIV 993
Qy 1227 IGLAKAQLIQIPFFPRNLNLTLLGLLHGLVFLPVILSVYG-PDVNPALALEQKAREANA 1285
Db 994 LCFARSEIFVVVYFQMYLALVIIGLHGLVFLVILSLAGPQLN--LDIEOQQTDEASS 1051
Qy 1286 AVN 1288
Db 1052 SLL 1054
RESULT 3
Ti5961
hypoetical protein F02B8.6 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: Ti5961
R:Miller, N.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F02B8.
A:Reference number: Z18436
A:Accession: Ti5961
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1456 <MIL>
A:Cross-references: UNIPROT:Q19127; UNIPARC:UPI000017B8AD; EMBL:U53340; NID:gl255859; P:
A:Experimental source: strain Bristol N2; clone F02B8
C:Genetics:
A:Gene: CESP:F02B8.6
A:Map position: X
A:Introns: 71/1, 124/3, 218/3, 269/3, 310/3, 403/2, 484/3, 528/2, 764/1, 858/3, 888/3, 1
Query Match 19.5%; Score 1347; DB 2; Length 1456;
Best Local Similarity 27.5%; Pred. No. 7,7e-86;
Matches 372; Conservative 258; Mismatches 565; Indels 158; Gaps 41;
Qy 76 KICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSNOSL 135
Db 63 EFCPLLHLDGN-KJCCTPSQABGLTKQIAQARHLIGRCPCDFNFKACMCEFTCSNQD 121
Qy 136 FINVTRVAQL--GAGQLPAVAYEAF-----YQHS--FAEQSYDSCSRVRVPAATAAVG 186
Db 122 FVSIEMKPIEKKEGFTPEYQPAEAVVNTVEYLSLTDFAEGHFSCKQVTFGQPALRV- 180
Qy 187 TWGVYGSALCAQRLNFGQDTGNGL-APLDITFHLLEPQAVSGGIQIP-LNEGVARCN 244
Db 181 -MC---TSTPCTLTNMLBFIGTQNLDLNIPHTKFLLYDPIKTPPSDRSTYMNVNFTGCD 236
Qy 245 ESQGDVATCSCQDCAACPAIARQALDSTFVLMQ-----PGSLVLIILCSVFAVVT 299
Db 237 KSARVGNPACSTSEC--NKEBYANLIDDDGTSQTCNVHGIACLNIPVIMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARDKSKVDPKKGTSLSKLSFSTHTLLGQPF 341
Db 295 LLICGVFTSYDEDTYNLRQTSQGESPKENIK-----RTGAWI-----HNPENNA 342
Qy 342 QGWGTWVASWPLTILVLSVIVPVALAAGLVFTLTDPVELWAPNSQARSEKAFHDQHP 401
Db 343 RDIGNMAGRNPKSHFFIGCAVLIFCLPGMIYHKESTNVVDMWSSPRSRARQEBWENANF 402
Qy 402 GPFPTNQVILTAPNRSSVRYDSLGLGPNFSGILDLDLLLELLELOERLRLHQLWMSPEA 461
Db 403 GRPQRYQQIMLL--SHRDFQSGKLYGP-----VFHKOIFEELFDILNAIKNISTQDSG 455
Qy 462 QNLSIQDICYAPLNPDTNSLYDCCINSLLOYFQNNRTLLLTANOT-----L 509
Db 456 -RTIILDDVCYRPMGPG-----YDCLIMSPFTNYFQNGKEHLDKMSKKEETVSEDDADPYF 510
Qy 510 MQTQSQVDWKDHFYCANAPLTFKDGFTALALSCMADYGAPEVFPFLAIGYKGYKDYSEARA 569

Db	511	SSEATTDEWNNHMAACIDQPSQK---TKSGLSGCMGTGCGPSAPNM-VFGKNSNTHQANS	567
Qy	570	LIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRTT	629
Db	568	IMWTILVQ--RTEBIEIKAELEKEFLKFCCKEYREKSPKI-PSFMAERSITDEIENDA	624
Qy	630	ABDLPIFATSYIVFLYISLALGSY-----SSMSRVMDVSKATGLGGLGVAVLGAVMAAM	684
Db	625	KDEIVTVIALAFLGIVYTFSLGRYFVCENQLWS-ILVHSRICLGLMSVILNLSFCSW	683
Qy	685	GFPSYLGIRSSLVILQVVPFLVLSYGADNIIFFVLEY--QRLP---RRPGEPREVHIGRA	739
Db	684	GIFSMEGHPVKNALVQVFFVTLGVCRTFMVKYAAQRVSMPEYKSPDQCPEI-VGMV	742
Qy	740	LGRVAPSMLLCSLSRAICFFIGALPMPAPVTFALTSGLAIVLDFLQMSFAVALLSLS	799
Db	743	MAGTMPAMFSSSLGCAFSFFIGGFTDLPAIRTCFLAGLAVIDLVLTCTFLALFVMDT	802
Qy	800	KQOEASRLDVCCCKPQELPPGQEGEGL-LLG-----FFQKAYAPFLHMI	844
Db	803	QRELG-----KP-BFFPYQIKDLGLGAYLIGRQATDTFTQTFHFQVAPFLMRM	853
Qy	845	TRGVLLFLALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYPT	904
Db	854	TRIIITFIASFIITVILSSKISVGFQDSMAFTEKSYISTHRYLDKFEVDGPPVPTV	913
Qy	905	TLYGNFSSBAGNAICSSAGCNFNFTQKIQVATEFPQSYLAIPASSWDDFDLWT-P	963
Db	914	DGELDWHRPDVQNKFTCPGCSDTSGNIMYAVGHTQYLSGEMYNWIDNYLEWISRK	973
Qy	964	SSCCRLYISGPNKDFCPSVNSL-----NCLNCH-----SITMGSV---RPSVEOFH	1009
Db	974	SPCCKYYVHDPN--TFCSTNRNKSALDDKACRTCMDFDVANSYPKSSIMYHRPSIEVY	1031
Qy	1010	KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASFMAYHKPL--KNSQDYTEALR	1067
Db	1032	RHLRHFLEDTNSECVFGGRASFKDAISFTSRGRIQASQFMTHFKKLSINSSDFIKAMD	1091
Qy	1068	AARELAANTADLRKVPGTDPAPFVPPYITITVFFEOYLTILPEGLFMLSCLVPTFAVS	1127
Db	1092	TARWVSRRLETSI-----DDTAHVPAYSKIFPFYEQYSTIMPILTTQLFITVGVGFI	1145
Qy	1128	CLLGLGLDRLSGLNLLSILVMILVDTVGFNALMDISNAVSLNLSVAGMSVEFVSHITR	1187
Db	1146	CVTLGIDVKGACAVTQVNSVFHIVAFWFIPIPNVLSATNLWSSGILLESVNVLK	1205
Qy	1188	SFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIIFPRNLNLT	1247
Db	1206	GYACSLRQAKORAESTVSGSIGPILSGPWTMAGSTMFLSGAHLQIITVYFFKLFLITI	1265
Qy	1248	LLGLLHGLVPLVILSYGPDVNPALALQKRAEAVAAVWVASCNHPHSRVSTADNIYV	1307
Db	1266	VSSAVHALITLPIALLAFGSGRSGSSETSTNDNEQHDACVLS--PTAESHSINVEEGIL	1323
Qy	1308	N-----HSFEGSIKGAGAINSEFLPNNGRQF	1332
Db	1324	NRPSLLDASHILDFLLKAEGLDKAI---GRDF	1353
RESULT 4			
S52525			
Probable membrane protein YPL006w - yeast (Saccharomyces cerevisiae)			
N;Alternate names: hypothetical protein Lpallw; hypothetical protein YP8132.07			
C;Species: Saccharomyces cerevisiae			
C;Date: 08-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004			
C;Accession: S52525; S59687			
R;Badcock, K.; Churcher, C.			
submitted to the EMBL Data Library, February 1995			
A;Reference number: S52519			
A;Accession: S52525			
A;Molecule type: DNA			
A;Residues: 1-1170 <BAD>			
A;Cross-references: UNIPROT:Q12200; UNIPARC:UPI000006A19C; EMBL:Z48483; NID:9683777; PID			

A;Experimental source: strain AB972			
R;Hall, J.; Ahmed, A.; Bussey, H.; Fortin, N.; Friesen, J.D.; Storms, R.K.; Vo, D.H.; Wa			
submitted to the EMBL Data Library, August 1995			
A;Description: The sequence of Saccharomyces cerevisiae chromosome XVI left arm.			
A;Reference number: S59677			
A;Accession: S59687			
A;Molecule type: DNA			
A;Residues: 1-1170 <HAL>			
A;Cross-references: UNIPARC:UPI000006A19C; EMBL:U33335; NID:9965087; MIPS:Y			
C;Genetics:			
A;Gene: SGD:NCRI			
A;Cross-references: SGD:S0005927; MIPS:YPL006w			
A;Map position: 16L			
C;Keywords: transmembrane protein			
F;1-17/Domain: transmembrane #status predicted <TM1>			
F;258-274/Domain: transmembrane #status predicted <TM2>			
F;341-357/Domain: transmembrane #status predicted <TM3>			
F;585-601/Domain: transmembrane #status predicted <TM4>			
F;614-630/Domain: transmembrane #status predicted <TM5>			
F;667-683/Domain: transmembrane #status predicted <TM6>			
F;698-714/Domain: transmembrane #status predicted <TM7>			
F;751-767/Domain: transmembrane #status predicted <TM8>			
F;1004-1020/Domain: transmembrane #status predicted <TM9>			
F;1027-1043/Domain: transmembrane #status predicted <TM10>			
F;1051-1067/Domain: transmembrane #status predicted <TM11>			
F;1103-1119/Domain: transmembrane #status predicted <TM12>			
F;1137-1153/Domain: transmembrane #status predicted <TM13>			
Query Match 19.2%; Score 1329; DB 2; Length 1170;			
Best Local Similarity 28.8%; Pred. No. 1.1e-84;			
Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;			
Qy	10	LLWAL-----LLRLAQSEPVYTHIQPGYCAFYDCGNP-----ELSGSLMTLSNVCSLN	60
Db	3	VLMTIALVGMLRLVQG-----TATCAMYGNCKKSVFGNELPCVPSPSPPPVLS	54
Qy	61	TPARKITGDHLLILOKICPLRYTGPNTQACSAQKLVLSLEASITKALLTRCPCASDNF	120
Db	55	ETSK-----LLVEVCGEWEKVR-YACCTKQDVVALRDNLQAKOPLISCCPACLNK	105
Qy	121	VNLCHNTCSPNQSLFINVTVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA	180
Db	106	NNLFCHFTCAADQGRFVNITKV-EKSKEDKDIIVAELOVFMNSSWASEFYDSCKNIKFSAT	164
Qy	181	ATLAVGTCMGVYGSAALNAQRWLNFGQDTGN--GLAPLDITFHLLPEQAVGSGIQPLNE	238
Db	165	N5----GYAMDILGGKKNYSQFLAKGLDAPKMLGSSPQINVKYDLANEE--KEWQEFND	218
Qy	239	GVARCNESQDDVATCSQDCAACPATARQALDSTTYLGOMPG---SLVLIIILCSVF	295
Db	219	EYVACDDAQ---YKACSDCQESCPHL-KP-LKDGVCVKGPLPCFSLVLIFFYFICALF	272
Qy	296	AVVTILLVGFVRVAPARDKSKVDPKGTSLSDK-----LSPSTHT-----LIGQFFQGWG	345
Db	273	AFPMYYLCKRKXKNGAMIVDDDIVPESG-SLDESETNVFESFNNEFNFGKLANLFTKVG	331
Qy	346	TWVASWPLTILVLSVIPVVALAAGLV-FTELTTDPVELWSAPNSQARSEKAFHDQHFQF	404
Db	332	QFSVENPKLITTVFSIFVFSFIIFYATLETDFINLWVSKNSKSEKFEKDFDDNFGFP	391
Qy	405	FRTNQVILTAPNRSRYRDSL--LLGPKNFSGILDLLELLELLEQLERLHLQVMSPEAQ	462
Db	392	YRTEQIFVNVNETGPVLSYETLHWFDVFN-----ITEEL-----QSS	429
Qy	463	RNISQDLCYAPLNPDNTSLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVQWKDHF	522
Db	430	ENIGYQDLCFRP-TEDST-----CVIESFYQYQG-----ALPNKDSWKREL	470
Qy	523	LYCANAPITFKDGTALALSCMADYGAIPVFPFLAIGYKGYKDYSEAEALIMTSLNYPAG	582
Db	471	QSCGKFP-----VNCPLTFQOPLKTNLL---FSDDDLNNAHAFVVVLTLLTNH---	514
Qy	583	DFRLAQAKLWEAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRTTASDLPIFATSYIV	642

Db 515 ----TQANWEER--LEEYLLDUKPEGL--RISFNTIEISLEKLN--NDISTVAISYLM 567
Qy 643 IFLYISLALSGSYSSMRVMDSKATIGLGGVAVVLGVMAAMGFFSYLGIRSLVTLQV 702
Db 568 MFLYATWALRRKDGKTRLL-----LGISGLLIVLASIVCAAGFLTLFGKSTLIIEAVI 621
Qy 703 PFLVLVSGADNIFIVLEYQV--LPRPGEPREVHIGRALGRVAPSMMLCSLSEAICFFIG 761
Db 622 PFLILAIGDINIFLITHEIDRNCQPEYSIDQKIISAIGRMSPSILMSLLCQTGCFLIA 681
Qy 762 ALTPMPAVTEALTSLGAVLDFLQMSAFVALLSLDSKQESRLDVCCKVPQELPPP 821
Db 682 AFTWPAVNFAITSVIFNGVLQLTAYVLSLSYKRSYKQIT----- 728
Qy 822 GOGEGILLGFFKAYAPFLHMTITRGVILLFLALFGVSLYSWCHISVGLDQELAPKDS 881
Db 729 -GNEETKESFLTFYFKMLTQ--KELIIIIISAFWFTSLVFLPELOFGLDQTLAVPQDS 784
Qy 882 YLLDPLFLNRYFEVGAPVYVFTTLGYNFSSBAGMAICSS--AGCNPFSTQKIQVATRF 940
Db 785 YLVDFPKDYSLFNLVGPVYVMVK--NLDLTKRQNOQKICGKFTTCERDSLANVLE---Q 840
Qy 941 PQSYLAIPASSWDDPIDWLTPSS--CCELIYSGNPKDKFCPSVNSLNCNCSITM 998
Db 841 RHRSTITEPLANWLDYFMLPQNDQCCRL---KKGTDVPCPPSPSRRC--ETCFQ--Q 894
Qy 999 GSVR-----PSVEOFHKVLPFLNDRPNKCPKGLAAYSVNLSDGQVLASRFMAY 1052
Db 895 GSWNNMGSFPFGKDFMEVLSWIN--AFSDPFLGGRAPYSTAL--VYNETSASVFRTA 952
Qy 1053 HKPLKNSQDYTEALRAARELANIATDLKKVPTDPAFVFPYTTNVPYQYLTLPRG 1112
Db 953 HPLRSQKDFIQAY-----SDGVRISSSPPELDMFAYSPPYIFVQYQTLGLPLT 1001
Qy 1113 LFWLSICLVPTFAVSCLLGLDRLSGLNLLSIVMLVDTVGFMAWDISYNAVSLINLV 1172
Db 1002 LKLGSAIILIFISSVFL--QNRSSFLALVVTMIIVDIGMALLGISLNAVSLVNI 1060
Qy 1173 SAVGMSVEVSHITSFAL---STPMTLERAKEATISMGSAVFAGVAMTNLPGLVLGL 1229
Db 1061 ICVGLGVECHVHSFVTPVETPKDANSRVLSINTIGESVIGKILTKFVGVCLAP 1120
Qy 1230 AKAQIQLIPFFRLNLLITLGLLHGLVFLPVLSYVG 1266
Db 1121 AQSKIPDFVFPFPMFTLIIVAAHALLFLPALISLFG 1157

RESULT 5
T30172
transmembrane protein patched - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30172
R:Goodrich, L.V.; Johnson, R.L.; Milenkovic, L.; McMahon, J.A.; Scott, M.P.
Genes Dev. 10, 301-312, 1996
A:Title: Conservation of the hedgehog/patched signaling pathway from flies to mice: indu
A:Reference number: Z20753; MUID:96176226; PMID:8595881
A:Accession: T30172
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1434 <GO>
A:Cross-references: UNIPROT:Q61115; UNIPARC:UPI0000021C02; EMBL:U46155; NID:g1181984; PI
A:Experimental source: clone M2, M9
C:Superfamily: Drosophila membrane protein patched

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 9e-37;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
Qy 356 LVLSVLPVVALAAGLVFTLTDDPVELWSAPNSQARSEKAFHQHGFPPFRN--QVILTA 414
Db 88 LVVGLLIFGAFVAGLKAANLETNVELVWEVGRVSRLEINTYTRQKIGBEAMFNQLMIQT 147

Qy 415 PNRSYRYSDDLGLPKNFSGIILDLLELLE--LQERLRLHQLQVMSPEAQRNISLQIDIC 472
Db 148 PKEEG-----ANVLTRALLQHLDSALQASRVHVVMYN---RQWKLEHLCY 190
Qy 473 AP-----LNPNTSLYDCCINSILQYFONNRTLLLLLTANQTLMGQ----- 512
Db 191 KSGELITETGYMDQIIEYLYPCLIIITPLDCPWEKALQSGTA--YLLGKPLRLWTNFDPL 248
Qy 513 -----TSQVMDKDHFLY-----CAN-----APL----- 530
Db 249 EFLBELKKNINYQVDSWEEMLNKAEGVGHGMDRPNCLNPPDPCPAPAPNKNSTKPLDVALV 308
Qy 531 -----TFKQGT-----ALALSCMADYGAPVPFPLAIGGYKKG 562
Db 309 LNGGCGGLSRKYMHWQEBELIVGTVKNAATGKLVSAHALQTMFLQMTPKQMYEHFQY--- 365
Qy 563 DYSEBALIMTFSILNYPAGDPLRLAQAKLWBAFLBEHRAPFORRMAGMFOVTTFAERSLE 622
Db 366 DY-----VSHINWNE---DRAAAAILSAWQRTYVEVWHQSVAPNSTQKVLPTT--TTLD 414
Qy 623 DEINRTAEDLPIFATSYIVIFLYISLALGSYSYSSRVMDSKATIGLGGVAVLGAVNA 682
Db 415 DILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKQGAAGVLAGVLLVALSVAA 469
Qy 683 AMGFFSYLGIRSSVLQVVPFLVSVGADNIFIVLEYQRLPRRPGPREVHIGRALGR 742
Db 470 GLGLCSLIGISFNAATTVQLPFLALGVGDDVFLLAHAFSETGQNKRIIPEDRTGCLXR 529
Qy 743 VAPSMMLCSLSEAICFFIGALTPMAVTRTFALTSLGLVLDLFLQMSAFVALLSLDSKQ 802
Db 530 TGASVALTISINVTAFPMALTIPIPALRAFSLQAAVVVFNFAVLLIIPAILSMOLYR 589
Qy 803 BASRLDVCCK-----VKQEL-----PPCQGGEL----- 827
Db 590 EDRLDIFCCFTSPCVSRVIVQEPQAYTEPHSNTRYSPPPYTSFSAHETHITWQSTVQ 649
Qy 828 ----- 827
Db 650 LRTEYDPHTHYTTTAEPRSEISVQPVTVTQDNLSQSPESSTSDRLLSQFSDSLHCL 709
Qy 828 -----LLGFPQAYAPFLHMTITRGVILLFLALFGVSLYSWCHISVGLDQELAPK 879
Db 710 EPPCTKWTLSFPAEKHYAPFLKPKAKVAVVILLFLGLLGVSLYGTTRVRDGLDLDIVR 769
Qy 880 DSYLLDYFLNRYFEVGAPVYVFTTLGYNFSSBAGMAICSSAGCNPFSTQKIQYAT- 938
Db 770 ETRYDFIAAQFKYFSP-----YNM-----YIVTKADYPNI 801
Qy 939 -----EPPEQSYLAIPASS-----WVDDFIDWL-----TPSS--- 965
Db 802 QHLLYDLHKSFNVKYVMLEENKQLPQMWLHYFRDLWLOQADAPDSDMETGRIMPNNYKN 861
Qy 966 -----CCRLYISGNKDKFCPSVNSLNCNCSITMGSVRSVRSVQFHKYL--PWFLN 1017
Db 862 GSDGGLVAYKLIVQGRDK--PIDISQIT--KQRLVDADGDIINPSA--FYIYLTAWSN 915
Qy 1018 D-----RPNKCPKGGAAAY--STSVNLSTSDGQVLASRFMAVHKPKNSOD 1061
Db 916 DPVAAASQANIRPHRPBWVHDK---ADYMPETRURIPAAEPYEAQFPFYLINGLRDTS 972
Qy 1062 YTEALRAARELANITA--DLARKVPTDPAFVFPYTTNVPYQYLTLPEGLFMLSCL 1120
Db 973 FVEALEKRVICNNYTSIGLSPNG-----YFP-----LFWEQYISLRHLLLSISVVL 1022
Qy 1121 VPTFAVSCLLGLDRLSGLNLLSIVMLVDTVGFMAWDISYNAVSLINLVASVGMVSVE 1180
Db 1023 ACTFLVCAVFLINPWTAGII--VMVALMTVELFGMGLIGIKLSAVPVVILLIASVIGVE 1081
Qy 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTNLPGLVGLAKAQLIQIF 1238
Db 1082 FTVHVALAFLTAGIDKN---HRAMLALBHMFAVLDDG--AVSTLLGLVLMAGSEDFIVRY 1137

QY 1239 FFRNLITLGLLGLVPLVILSYVG--PDVNPALALEQ---KRAEEAVALVAVASCP 1293
DB 1138 FFAVLAILTVLGLVGLVLLVLLSFFGCPCEVSPANGNLRLPTSPSEPPSVVRFAVPP 1197
QY 1294 NHPRSVS--TADNIYVNHSEFSGI-----KGAG 1319
DB 1198 GHTNNGSDSSSEYSSQTTVSGISEELRQYEAQQGAG 1234
RESULT 6
T18538
patched protein - chicken
C:Species: Gallus gallus (chicken)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T18538
R:Marigo, V.; Scott, M.P.; Johnson, R.L.; Goodrich, L.V.; Tabin, C.
Development 122, 1225-1233, 1996
A:Title: Conservation in hedgehog signaling: induction of a chicken patched homolog by S
A:Reference number: Z18958; PMID:96205046; PMID:8620849
A:Accession: T18538
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1442 <MAR>
A:Cross-references: UNIPROT:Q90693; UNIPARC:UPI00001328B4; EMBL:U40074; NID:G1335850; P1
C:Gene: PTC
C:Superfamily: Drosophila membrane protein patched
C:Keywords: transmembrane protein
Query Match 9.1%; Score 628.5; DB 2; Length 1442;
Best Local Similarity 23.6%; Pred. No. 1.6e-35;
Matches 284; Conservative 165; Mismatches 421; Indels 335; Gaps 47;
QY 365 ALAAGLVFTTLDTPVELMSAPNSQARSEKAFDQHFQFFFTN-QVILTAPNRSRYD 423
DB 111 AFAGVGLRAANLSTNVEELMVEGVRSRELNTRQKIGEEAMFNQPMIQTPOEDG---- 166
QY 424 SLLGPKNFGIIDLLELLE--LQERLRLQWSPQAQRNLSLQDICYAP----- 474
DB 167 -----TNVLTTEALRQHLDSALQASRVHYMYN-----RQWKLEHLQYKSGBLITEA 213
QY 475 --LNPDTSLYDCINSLLQYFQNNRTLLLTANQTLMGQ----- 512
DB 214 GYMDQIIEYLYPCLITPLDCFWEGAKLQSGTA--YLLGKPLQWNPDLFLELKKI 271
QY 513 TQVDMDKHFLY-----CAN-----APLTF--KQGT-----ALALS--CMADY 546
DB 272 NYQVESWEEMLNKAEVGHGYMDRCPCLNPADPCPITAPNKNSTKPLDVALVLSGGC---Y 328
QY 547 GAPVFFP-----LAIGG-----YKGDYSEALIMTFSL-----NNYP----- 580
DB 329 GLSRKYMHWQELIIGTVKNSGKLVS--AQAQTMTFQMTKQMYEHFKGYEYSHINW 387
QY 581 AGDPLRLAQLMEAEFLEEMRAQRMRMAGMFQVTFTEASRLEDEINRTTAEDLPFATSY 640
DB 388 NEDKAAALEAWQRMVEVHVHQSVAQNSQKVLSTFT--ITLDDILKSPSDSVIRVASGY 446
QY 641 IVTFPLYSIALGYSNWSRVMDSKATLGLGVAVVLGAVMAAMGFYSYLGRSSVLILQ 700
DB 447 LMLAVACLTM---LRWD---CAKSGOAVGLAGVLLVALSVAAAGLGLCSLIGISFNAATQ 501
QY 701 VVPLVLSVGCADNIFIVLEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEACFFL 760
DB 502 VLPFLALGVGDVDFLLAHAFSTGQNKRIPFEDRTGECUKRTGASVALTSISNVTAFPM 561
QY 761 GALTMPAVRTFALTSGVLVILDLQMSAFALLSLDSKQREASRLDVCCC----- 812
DB 562 AALIPALRAFSLQAQVAVVVFAMVLLIFPALLSMDLVRRDRDLDFCCFTSPCVTR 621
QY 813 ---VKPO-----ELPPQGGEGE----- 827
DB 622 VIQIEPOAYANDNICYSSPPPYSSSHFAHETQITWQSTVOLRTEYDPHTQAYYTTAEPR 681

QY 828 -----LLGFQKAYAP 838
DB 682 SEISVQPVTVTQDSLSCQSPESASSTRDLLSQFSDSSVHCLPECTKWTLSSTFAEKHYAP 741
QY 839 FLHWITRGVLLFLALFVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 898
DB 742 FLKPKAKVVIIFLGLGLLSLYGTRVRDGLDLTDIVPRDREYDFIAAQPKYFSF-- 799
QY 899 PVYFTVLGYNFSSEAGMNAICSSAGCNFSPFKIQYAT-----EPPEQSYLAI 948
DB 800 -----YNN-----YIVTKADYPNVQHLHYELHRSFSNVTVLL 833
QY 949 PA-----SSWDDDFIDWL-----TFSS-----CCRLYISGPNKD 977
DB 834 EGRQLPKWHLFRDWLQGLQDAPDSDWETGKITTSYKNGSGDDAVLAYKLLVQTGNRA 893
QY 978 KFCPSTVNSLNCNKNCSITMGSVRSVPEQFHKL--PWFLND-----RPNI 1022
DB 894 K--PIDISQLT--KQRLVDADGIIINFNA--FYILTAWVSNDPVVAYAAQANIRPHPEW 947
QY 1023 KCPKGGGLAAV--STSVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELAANITA--D 1079
DB 948 VHDK---ADWMPETRLRIPAREPIEYAOFFPYLNGLRSETSDFVEAIEKVRACNNYISLG 1004
QY 1080 LRKVGPTDPAFEPVPPYITITNVFYEQYLTILPEGLFMLSCLCVPTFAVSCLLGLDLRSGL 1139
DB 1005 IASYPNG-----YFP-----LFWEQYIGLRHLLLSISVVLACTFLVCALFLNPNWTAGI 1054
QY 1140 LNLISVIMLVDTVGFMAWLDISYNAVSLINLSAVGMSVEFVSHITRSP--AISTKPTW 1197
DB 1055 I-VVVLAMTVLVELFGMGLIGIKLSAVPVVILIASVGIVGEFTVHIALAFUTALGDKN-- 1111
QY 1198 LERAKEATISGSAVFAVAGVAMTNLPGLVLGLCAKAOIQIFFFFRLNLIITLLGLLHGLVF 1257
DB 1112 -RRVLALEMFAPVLG-AVSTLLGVMLAGSEDFIVRYFAVLAITLLTGLVGLVLT 1169
QY 1258 LPVILSYVG--PDVNPALA---LEQKRAEAAVAAVMVASCNHPSRSTADNIYVNHSE 1312
DB 1170 LPVLLSPFGPYPEVSPACGRNRLPTSPPEPPPSIVRFALPFGHTNNGS--DSSDSEYSSQ 1227
QY 1313 GSIKG 1317
DB 1228 TVVSG 1232
RESULT 7
S44797
F09G8.4 protein - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 09-Jul-2004
C:Accession: S44797
R:Anderson, K.
submitted to the EMBL Data Library, February 1993
A:Description: Sequence of the C. elegans cosmid F09G8.
A:Reference number: S44797
A:Accession: S44797
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-915 <AND>
A:Cross-references: UNIPROT:P34389; UNIPARC:UPI000013B859; EMBL:L11247; NID:G156280; P1D
C:Gene: C
A:Introns: 51/1; 214/3; 255/3; 326/3; 382/2; 409/3; 455/3; 505/3; 569/2; 702/3; 836/3
Query Match 8.7%; Score 600; DB 2; Length 915;
Best Local Similarity 22.7%; Pred. No. 8.9e-34;
Matches 236; Conservative 158; Mismatches 422; Indels 224; Gaps 37;
QY 10 LWMALLRLAQSEPTTHIQPGYCAFYDECGKNPLSGSLMTLSNVSCLSNTPARKITGD 69
DB 14 VLFLLLHLALCQ-----AKVM--TECDGEEDSNHPPCKTNKSTYLPITVTRSLNPT 64
QY 70 HLILLOKICPLRYTGN--TQACCSAKOLVLEASISITKALLTRCPACSDNFVNLHCHNT 128

Db

643

-----DDNGTMVKSFRFILGMKDLVTMTDQTDATATSFREVAARM-----

681

Qy

1081

RKVPCTDPAFVPPYTTINVEYQVLTILPGLFMLSCLVPTTAVSCLLLGLDLRSGLL

1140

Db

682

-----PEFNVTTFMPIMETDQYIIIPNTQNIILALVMIVAVLFIPQPMCS-LW

733

Qy

1141

NILSIVMLVDTVGPMALWDISYNAVSLINLVSAVGMSEVFSVHITRSFAISTKPTWLER

1200

Db

734

VALACASIDFGVIGTWTUWNLWDIAISMITIIMSIGFSDVSAHAYGVVSRDITAAGR

793

Qy

1201

AKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFRLNLITLGLLHGLVFLPV

1260

Db

794

VKEALSALGWPLSQG-ANSTIIAVSVLADIPAYMI-VTFEKTVVLISLGLLHGLVFLPV

851

Qy

1261

ILSYV-----GPDVNPALALEOKRAEBA-----VAAMVAVSCPNHPSRVS

1300

Db

852

LLSIFVRGCCIIIPSPHPGHPSAQKIEQIRIAATSSSPLDLRTVAPLRASSPIFPFPHRLE

911

Qy

1301

-TADNIYVNHSPGSIK 1316

Db

912

YDESPTVHNRSKNSIK 928

RESULT 9

T13952

membrane protein ptch2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: T13952

R:Motoyama, J.; Takabatake, T.; Takeshima, K.; Hui, C.

Nature Genet. 18, 104-106, 1998

A:Title: Ptch2, a second mouse patched gene is co-expressed with Sonic hedgehog.

A:Reference number: Z17830; MUID:98122566; PMID:9462734

A:Accession: T13952

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1182 <MOT>

A:Cross-references: UNIPROT:O35595; UNIPARC:UPI0000001587; EMBL:AB010833; PIDN:BAA24691.

A:Experimental source: strain BALB/cCrSlc

C:Gene: ptch2

C:Superfamily: Drosophila membrane protein patched

C:Keywords: transmembrane protein

Query Match

Best Local Similarity

Matches

282;

Conservative

154;

Mismatches

433;

Indels

297;

Gaps

44;

Qy

330

SFSTHTLLGQ-----FFQG-----WGTWVASWPLTILVLSVIPVVALAAGLVFTELT

376

Db

20

SSAPHILAGSLQAPLMIRAYFQGLLSLCRIQKCHGVFLGLVAFGALGLRVAVIE

79

Qy

377

TDPELWSPNSQARSEKAFDQHG-PPFRNTQVILTAPNRSSVRYSDLLGPKNFSGI

435

Db

80

TDLEQLWVEGSRVSQELHYTKELGEEAAYTSQMLI---QTAHQEGGNVLTPE-----A

131

Qy

436

LDLDLLELLELQERLRLHQVMSPEAORNISLQDICY---APLNPDN-----TSLYDCC

486

Db

132

LD-----LHQAALTASKVQVSLYKGSWDLNKICYKGVGPIENGMIERIEKULFPVC

184

Qy

487

INSILQYFQNNRTLLLTANQTLMGQTSQVDWKKD---HFLYCANAPLT-----FKQGTALA

539

Db

185

ILTPDLCFWGAK---LQGSAYLPGRPDIQWNTLDPQLLELGPFPASLEGFRELDKA

241

Qy

540

LSCMADYGAPVF-----PFLA---IGGYKQKQYS-----

565

Db

242

QVGQAYVGRFCLDPDDPHCPSPAPNRHSRQAPNVAVQELSGGCHGFSHKFMHMOEILLGG

301

Qy

566

-----EAEALINTF-----SLNYPAGDPRLAQAKLWEAEFLMEMRAFORRMAG

609

Db

302

TARDLOQLRAEALQSTFLLMSPROQYEHFRGDYQTHDIGNSEEQASWVLQAWRRFVQ

361

Qy

610

MFQVTFTAERSLE-----DEINRTTAEDLPFATSYIV--IFLYISLALGSSWS

658

Db

362

LAQEALPANASQOIHAFFSSTTLDDILRAFSE-----VSTTRVVGGLMLLMAACVMTLWRWD

417

Qy

659

RVWVDSKATIGLGGAVALVGLGAVMAAGFFSYLGRSLVILQVVPVVLVLSVGDADNIFIV

718

Db

418

--CAQSGAVGLAGVLLVALAVASGLGCLGALGITNAATTQVLPFALGIGVDDIFLLA

475

Qy

719

LEYQRLPRRPEGEPREVHIGRALGRVAPSMMLCSLSBAICFFLGAALTPMPAVRTALTSGL

778

Db

476

HATKAP--PDTPLPERMGECLASTGTSTVALTSVNNNVAFMAALVPALPAFSLQAAI

533

Qy

779

AVILDFLLQMSAFVALLSLDSKQEAASRLDVCC-----VKPQE-----

817

Db

534

VVGCNFAAVMLVPAILSLDRHRQRDLVLCFCFSPCSAQVIQMLPQELGRDRAVPVGI

593

Qy

818

-----LPP-----PG-----QSEG--

826

Db

594

AHLTATVQAFTHCEASSQHVVTILPQAHLLSPASDPLGSELSPGGSTRDLLSQEGTG

653

Qy

827

-----LLLGFFQKAYAPFLHWTIRGWVLLFLALFGVSLYMCHSISVGLDQ

873

Db

654

POAACRELLCAHWTLAHFARYQAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL

713

Qy

874

ELALPKDSYLLDYFLNRYFEVCAVYFVTTLYGNFS-SEAGHNAICSSAGCNPSTQ

932

Db

714

TDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHSORAL-----FDLHQ

761

Qy

933

KIQVATEFPQSILAIPAS-----SWDDF-----IDMLTPSSCCRLYISGPNK

976

Db

762

RFSSL-----KAVLPPPATQAPRTWLHYRSWLOGIOAAFDQDWDASGRITCHSYRNGSD

816

Qy

977

D---KFCPSVTNSLNCIKMCSITM-----GSVRPSVQFHKYLPWFLNDRNRIKCPK

1026

Db

817

GALAYKLLIQTGNAQEPDLFSQLTTRKLVKDEGLIPP--ELFYMGLTVWVSSDPL-----

869

Qy

1027

GGTAA-----YST---SVNLTSQGVLASRFMAYHKPLKNSQDYTRALR

1067

Db

870

-GLAASQANFYPPPPPEWLHDKYDTTGNLRIPAAQPLEFAQFPFLHGLQKTAQDFVEAIE

928

Qy

1068

AARELAANI-TADLRKVPGTDPAPFVFPYTTINVEYQVLTILPGLFMLSIC--LVPTF

1124

Db

929

GARAACTEAGQAGVHAYPSGSPF-----LFWEQYGLG--RRCPFLAVCILVCTF

976

Qy

1125

AVSCILLGLDLRSGLNLLSIVMILVDTVGPMAWLDISYNAVSLINLVSAVGMSEVFSVH

1184

Db

977

LVCALLLSPTWAGLI-VLVLAAMTVELFGIMGFLGIKLSAIPVVLVASIGIGVEFTVH

1035

Qy

1185

ITRSPAISTKPTWLERAKEATISMGSAVFAGV---AMTNLPQILVLGLAKAQLIQIFFFR

1241

Db

1036

VALGELTSHGSRNLRAA-----SALEQTFAPVTDGAVSTLLGLMLAGSNFDFIIRYFV

1090

Qy

1242

LNMLTLLGLLHGLVFLPVILSYVGP 1267

Db

1091

VLTUTLLGLLHGLLVLPLVLSILGP 1116

RESULT 10

T36746

hypothetical protein Y39A1B.2 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26746

R:Wall, M.

submitted to the EMBL Data Library, January 1998

A:Reference number: Z20258

A:Accession: T26746

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1003 <WIL>

A:Cross-references: UNIPROT:O9XXR9; UNIPARC:UPI00000782FA; EMBL:AL021482; PIDN:CAAL6339

A:Experimental source: clone Y39A1B

C:Gene: CESP:Y39A1B.2

A,Map position: 3
A;Introns: 89/2; 126/3; 167/2; 194/2; 232/1; 266/3; 295/2; 323/2; 363/1; 519/3; 696/3; 7

	Query Match	8.3%; Score 573.5; DB 2; Length 1003;
	Best Local Similarity	22.8%; Pred. No. 7.3e-32;
	Matches	236; Conservative 198; Mismatches 376; Indels 225; Gaps 42;
Qy	339 QFFQGGTW---VASNPLTILVLSPVVALAAGLVFTE--LTTDPVELMAGPNSQARSE	393
Db	9 RPAHAFGAYGVVVRHPPFFIIPILITAGLSTGLLRHQAFMKDELELYTPTDAQARKE	68
Qy	394 KA-----FHQHFQPFRTNQ-----VLTAPNRSSVRYDSLLLPKPFSGILDLDL	440
Db	69 LSOLDHLFHINDSPFYATRYDIRRAGYIIVT-----NOBEDGDILNPLVMHSAQOLWS	123
Qy	441 LLELLELOQLRLHQLVMSPEAQRNISLQDIC-----YAPLNPNPTSLYDCCI	487
Db	124 IVQSLTVED-----EDDRINYPSICVKEPPIPEPFSKALHSLFAPNWTPEEICV	173
Qy	488 -NSLLQYFQNNRTLLLTANQTLMGOT-----SQVDMKDHFLYCANAPLTFKOGTALAL	540
Db	174 SNPLVEIFK-----LLLVSDRSFLNRSIDEMTLLSQI-----SDAIPQDSGWMTHL	218
Qy	541 SCWADYGAPVPPPLATGGY---KGKDYSEBALIMTFSLNYPAGDPRLAQAOKLWEBAFL	597
Db	219 -----LGGVTLDDDKRIAGAKAMLLPYALRH--SSDDSDWAEKWEVEFLA	261
Qy	598 EEMRAFQRRMAGMFQVTFEATERSLEDEINRTTAREDLPIFATSYIVIPILYISLALGSSYSW	657
Db	262 DFLLOYD---SPIRASWYTYETLAESAARDRIQLIHMLLPFCVCSIF-TIACCVSFW	317
Qy	658 SRMVDKATLGLGGVAVVLGVAAMGFFSYLGRSLVLQVVPFLVLVSGADNIPF	717
Db	318 RR---SRPWLAIQGVISAAMATASAVGILLLAGYGMTSVAYS-MPPIVPSVGVDNVFIL	372
Qy	718 VLEYQRLPRPGEPRVHIGRALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRFTALTS	777
Db	373 LSAMRSTST--ETLHRMKETPADAAVSTVTSLTDLISFGVGCATPPFSVQMFCAVAV	430
Qy	778 LAVILDFLLQMSAFVALLSLDSKROBASRLDCCCV-----KQBELPPQGQEG-----	826
Db	431 AAVITYIYQLTFFFAAVMVYTNREINR-----HCIFPHKJUKDTLPKIAAQGRDSPEK	486
Qy	827 --LLLGFFOKAYAPFLHMITRGVLLFLFALFGVSLYSMSCHISVGLDQELALPKDSYLL	884
Db	487 NTILAQFFRTTYSDFLLNPLVRIVILVTFCVYLGVASYGCTKVKLGLEPNLLPENSXGK	546
Qy	885 DYFLFLNRYF-EVGAPVYFVTTLGYNPFSS-EAGNNAICSSAGCNPFSTQKIQVATEPPE	942
Db	547 RTLMAAEKFSYDYGSSLH---VMYNLSYEDVAPRKIWN-----VLEKEVELYE	592
Qy	943 QSYLAI PASSWDDDFIDW-----LTPSSCCRLYISGPNKKFCPSVTNSLNCILKNCS	995
Db	593 HTEFTASSUSWLRTFLAFVQAGLLITPE-----NPFYILKN-----	629
Qy	996 ITMGVSRPVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGOVL-ASRFMAHYK	1054
Db	630 -----VFLSQPF-----AKYNRDVLVTEDGEHLEASRI-----	658
Qy	1055 PLK-----NSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYTTINVFYQVLTLP	1110
Db	659 PVQLRHVGSANOSRAMRLFRLAE--TSELQ---TGVIADFPQ-----FAEQVNAVLP	706
Qy	1111 EGLFWLSLCLPPTFAVSCLLGLDRLSGLNLLSIVMLVDTVGFMAWDLISYNAVSLIN	1170
Db	707 GTLSSIAYAGVAVVSLILIPFVAS-LWVSFISVINIGILGFMFTFWSVRLDPIFSMT	765
Qy	1171 LVSAGVMSVEFYSHIRSFASITKPTWLERAKKATISMGSAVPAGVAMTNLPGLVLGLA	1230
Db	766 IWSIGFCVDFAAHLAYNAFKQMGDSERBNALYAVGAPILMS-ATSTIIGVSPWASA	824
Qy	1231 KAQLIQIFPRLNLLITLGLLHGLVFLPVLIS---YVGPDVNPALALEQKAEBAVAV-	1287

Db 825 ESYVFR-SFLKTIIMLVILGALHGLVLPLLVSFPGCG-----SSKKAKEHIDAVD 875

Qy 1288 -MVASCNPHPRSVT 1301
 + : |:| | : :

Db 876 QKLQAQYNPAPTAS 890

RESULT 11
Tl8291
patched protein - zebra fish
C:Species: Brachydanio rerio (zebra fish)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: Tl8291
R:Concordet, J.P.; Lewis, K.E.; Moore, J.W.; Goodrich, L.V.; Johnson, R.L.; Scott, M.P.
Development 122, 2835-2846, 1996
A>Title: Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog signaling in zebrafish gastrulation
A:Reference number: Z18860; MUID:96379744; PMID:8787757
A:Accession: Tl8291
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1220 <CON>
A:Cross-references: UNIPROT:Q98864; UNIPARC:UPI00001328B3; EMEL:X98883; PIDN:CAA67386.1
C:Genetics:
A:Gene: ptcl
C:Superfamily: Drosophila membrane protein patched

Query Match 8.3%; Score 572; DB 2; Length 1220;
Best Local Similarity 22.4%; Pred. No. 1.2e-31;
Matches 270; Conservative 174; Mismatches 431; Indels 328; Gaps 43;

Qy 321 KGTSLSDKLGFSTHTLLGQPFGQGTVWASWPLTLVLNVSPVVVALAGLVFTTEITDPV 380
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||
Db 51 KGKAVGGCAKPLIRARFOAFSLSGCHIQRHCGKVLPFILGVFGALSVDGLRAAIETDIE 110
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 381 ELWSAPNSQARSEKAFDHQHFG-PFFRTNQVILTAPNRSSYRYDSLLLGPNFSGILDLD 439
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 111 KLWVEAGSRVSKELYTKRGQGESVTSOMLIQTPKQG-----TNLTQE 157
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 440 LLLELLEQLERLRHLQVWSPAQRNISLODI CYAPLNP--DNT-----SLYDCINSL 490
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 158 AL--LLHLEAALSASKVOVSLYGKSNDLNKLCFKSGVPPIENVMIERMIDKLFPCMI VTP 215
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 491 LOYPQNRTL-----LLLLTANQ---TLMGQTSQVDWKOHFLYCA 526
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 216 LDCEWEGSKLGGSAYLPGMPDIOWMNLDPKLMEELSQTFSLEGFREMLDKAQVGHYM 275
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 527 NAPLTFKDGTALALSCHADYCAP-----VPPFLAI-----GGYKG-----561
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 276 NRCLDPSDT----DC--PHSAPNKPDQVFNIAEELQQGCGHGFSKPFMWQESLILGER 329
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 562 -KD----YSBAELIIMTFSL-----NNYPAGDPRLAOAKL-----WEBAFLEE 599
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 330 VKDSNALQSABALQTMFLMSPKQLYEHFKDDTEIHINDWNEDKATALESWORKFPVEV 389
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 600 MRAP--ORRMAGMQPVTTFAERSLEDINKRTTABDLPIFATSIVIFYISIALGSYSW 657
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 390 VHGISPNQSSSNVVAFFT---TLNDIMKSFSDSVIRVAGGYLLMLLAYACVTM---LRW 443
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 658 SRVMWDKATILGGVAVVLGAVMAMGPSYIGIRSLLVILOVVPFLVLSVGADNIPIF 717
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||::~||::||:

Db 444 D--CAKSGQAGVGLAGVLLVALSVAAGLGLCSLLGSFNAAATTQVLP SLALGI GVDDMF-- 500
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 718 VLEYORLPRAPGEPREVHI GRALGRVAPSMLCLSBAICFFLGCALTMPPAVRTPALTSQ 777
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 501 -LGHSTFTTSNIPFKERTGDCLRTGTSTVALTSVNMMIAFFMAALVIPAPRAFSLQNA 559
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 778 LAVILDFLLQMSAFVALLSLDSKRQESRASRLDVCC-----VKPE-----817
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Db 560 VVVVFNFAMALLIPAILSLDLHERREDKRLDILCFYSPCSSRVIIQIQPQLSDANDNHQ 619
 ||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||:||||

Qy 818 -----LPPPQQ-----823
 |||||

Db 620 RAPATPTYTGSTITTTHTTTTTVOAFTQCDAAGHVITLIPFTSQISTPTPSMWLSTPT 679
 |||||

QY	824	-----GEGL-----LLGFFKAYAPFLLHWTR	846
Db	680	TTDPVGSQVTTSSSTRDLAQBEKREGRECVLPFFRNMNLSFAREKAPLLKPKETK	739
QY	847	GVVLLFLALFGVLSYNCHISVGLDQELAPKDSYLLDYFLNRYFEVGAPVFEVTL	906
Db	740	TVVVVFALLSLSLYGTWHDGLYLDIVPRDTQBEYFTAQFKYFSP-YNNMLVTMD	798
QY	907	GVNFSSEAGMNAICSSAGCNFFSTQKIQYATEPEQSILA-----IPASSWDDFDW	960
Db	799	GFDIAR-----SQQLQLHNAFNSVKYVVKDGNHKLIP-RMMLHYFQDW	841
QY	961	LTPSSCC-----RLYISGNPKKFPSTVNSLNCNKC	993
Db	842	LKGLQATFDADWEAGKITDYSYRNGTDCALAYKPLIQTGSKGPPFNSQLTSRRLVDG-	900
QY	994	MSITWGSVRPSVEQHKYLP-WFLNDRPNIKCPKGLAA-----YST---	1034
Db	901	-----DGLIPPEV--FYIYLTWVNSD-----PLGYAASQANFYHPREWHDKYDTGE	948
QY	1035	SVNLTSDGOVLASRFMAYHKPKNSQDYTEALRAARELAAN-ITADLRKVPCTDPAPEVF	1093
Db	949	NLRIPAAEPLEBAQPFYINGLRQASDFEAIESTICEBFMRQIKNYENPENG-----Y	1002
QY	1094	PYTITNVFEQYLTLPGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIVMLVDTV	1153
Db	1003	PF-----LFEQYIGLRHWFLLSISVVLACTFLVCAILLNPNMTAGVI-VFILPMWTVELF	1057
QY	1154	GFMAWDISYNAVSLINLSVAGMSVEFVSHITRSF--AISTKPTMLERAKEATISMGSA	1211
Db	1058	GIMGLIGIKLSAIPVILLASVIGVEFTVHIALGFLTAIGDRNT---RSAVAMEHMEAP	1114
QY	1212	VFAGVAMTLPGLVLGLAKAQLIQIFFRMLNLLTLGLLHGLVFLPVILSYGCP--DV	1269
Db	1115	VIDG-AISTLLGVMLAGSEFDFIMRYPPFAVLAITLLGILNGLVLLPVLLSLMGPPAEV	1173
QY	1270	NPA 1272	
Db	1174	VPA 1176	
RESULT 12			
T29590			
hypothetical protein F55F8.1 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C:Accession: T29590			
R:Gatting, S.; Scheet, P.; Kemp, K.			
submitted to the EMBL Data Library, November 1996			
A:Description: The sequence of C. elegans cosmid F55F8.			
A:Reference number: Z20647			
A:Accession: T29590			
A:Status: preliminary; translated from GB/EMBL/DDBJ			
A:Molecule type: DNA			
A:Residues: 1-889 <GAT>			
A:Cross-references: UNIPROT:P91346; UNIPARC:UPI000017BA20; EMBL:U80447; PIDN:AAB37812.1;			
A:Experimental source: strain Bristol N2; clone F55F8			
C:Genetics:			
A:Gene: CESP:F55F8.1			
A:Map position: 1			
A:Introns: 36/3; 66/3; 98/3; 149/2; 200/3; 240/2; 279/1; 358/1; 481/2; 569/3; 606/3; 697			
Query Match			
Best Local Similarity 7.8%; Score 539.5; DB 2; Length 889;			
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;			
QY	401	FGPFRFNQVIL-----TAPNRSSRYDLSLLGPKNFSGIILDLLLELLE	446
Db	31	FGPYSYTERRIIHDAWPLVDGTFVAGRAVTS--REVQVAVVARSGNILLDRVFSNELKL	88
QY	447	LOEERLRLQVMSPAQONISLQDICYAPLNPDTSLYDCCINSILQYFQNNRTLLLTAN	506

RESULT 13

T27969

hypothetical protein ZK675.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T27969

R:Sim, M.

submitted to the EMBL Data Library, November 1994

A:Reference number: Z20448

A:Accession: T27969

Db	89	MESFIRN-NITVQFSNRTWSFADLCIA--GPDGR-----CANNDHIQ-----	127
QY	507	QTLMGQTSQVDKQDHPFLYCANAPLTKDGTALALSCHWADYGAPVFPFLAIGGVK-	560
Db	128	--LASRLHQ-----HGINTYPTVRLSDKSAYIAS-----ALGGVKLAKGDN	167
QY	561	GKD-YSEAEALIMFTSLNNYP-----AGDPRLAQAKLMEEAELEEMRAFQRMAGMFQV	613
Db	168	GENIIVENTAWLLIYQLKFPNEISYVSG-----LWEREPKNMDEY-KKQAKYISI	218
QY	614	TFTAERLSDEINRTTAEDLPFATSYIVIFLY-----ISLALGSSSSSRVWVDSKAT	667
Db	219	TYFHSQTLSDLNRAERLAPKEIGAFVLVCFSLCSIVTIKSGVYDW-----VVTXPI	274
QY	668	LGLGGVAVVLGAVMAAGPFSYIGIRSSLVILQVWPFVLVSVGADNIFIFVLEYQRLPRR	727
Db	275	LSVLGVSNAAGMGLASAMGLTYLEIQN-DIIAMFPLVAVAGTDNMFVMSAKRTDR-	332
QY	728	PGEPREVHIGRALGRVAPSMMLCSLSEACFFLGALTPMPAVRTFALTSGLAIVLDFLLQ	787
Db	333	-NLKYDQRIAEACMAADAVALITALDALSFGVGTITTIIPAVQIFCIYTMCALLTFAYQ	391
QY	788	MSAFVALLSDSKRQE-----ASRLDV-----CCCVKPOELPPCQG----	824
Db	392	LTFPCAILVYTRIIEEQGLHSIWLRPATVYSTSPLNVKLFWLGSQPK-PLPSCGTVSS	450
QY	825	-----EGLLGPFQKAYAPFLH-WITRGVVLALLFLALFGVSLYS	863
Db	451	TSSVSTWTSQATSPASXHLHCAATSFRRWYAPVLMPQWI-RAIAGLWLYLIGISYG	509
QY	864	MCHISVGLDQELAPKDSYLLDYFLNRYFEVGAPVYFVTT-----LGYNSSE	913
Db	510	CTHLKEGLEPANLLVDDSYATPHYRVLEKHYWHYGASLQIVVSNPPDLRDPVERINNDRM	569
QY	914	AGMNAICSSA-GCNFSE-----TQKIQYATE-----FPEQSYLAIPASSWVDDFI	958
Db	570	ASTFANCKVAIGDSDVQFWLREMQVSEBIIHQYDNBEKFDYDHAQAQYIYSDMSQPVVDV	629
QY	959	DWLTSPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQHKYLPWFLND	1018
Db	630	-----WGRNN	634
QY	1019	RPNIKCPKGLAAYSTSVNLTSQGVLASRPMAYHKPKNSQDYTEALRAARELAANITA	1078
Db	635	-----NSERIIKTFRPMIGNRDISTTTKQTEATNTFREIASRP--	672
QY	1079	DIRKVPGTDPAFEPFYITNVFYEQYLTLPGLFMLSCLVPTFAVSCLLGLDLRSQ	1138
Db	673	-----EQYNVYTYMPLWLTDDQVALVVPNTMQDIIIVACMLVISALLIPOPVCSP	723
QY	1139	LINLLSIVMLVDTVGFMAWDISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWL	1198
Db	724	WV-AVTIGSIDLGVLFMTLNNVLDALSMITIMSVGFSVDYSAHITYAVVISKESTTS	782
QY	1199	ERAKEATISMGSAVAGVAMTNLPGLIVLGLAKAQLIQIFFRMLNLLTLGLLHGLVFL	1258
Db	783	ARVCDALGDLGMPVAQAG-AMSTILAVSLSDVPATMI-VTFEFTVFLAISIGFLHGLVFL	840
QY	1259	PVILS-YVG 1266	
Db	841	PLMLSVFVG 849	

A;Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1405 <MIL>
A:Cross-references: UNIPROT:Q09614; UNIPARC:UPI0000835A5; EMBL:Z46812; PIDN:CAA86843.1.L
A:Experimental source: clone ZK675
C:Genetics:
C:Gene: CESP:ZK675.1
A:Map position: 2
A:Introns: 31/1; 59/3; 100/3; 244/2; 485/1; 950/2; 1298/2; 1335/3
C:Superfamily: Drosophila membrane protein patched

Query Match 7.4%; Score 514; DB 2; Length 1405;
Best Local Similarity 20.2%; Pred. No. 1.7e-27;
Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

Qy 321 KGTSLDKLSFSTHTLLGQPFQCGWTW--ASWPLTILVLSVIPVVALAAGLVFTETLTDD 378
 :
Db 100 RGRATGNRYALYSRLQLKLPAFGNTVRHNSMI--ILAVSMIFAVC-CYGLQVVIETD 157
 :
Qy 379 PVELWSA-----PNSQARSEKAFHQHFQPPF-RTN-----QVILTAPNRSS 419
 :
Db 158 IVKLWAQGGRLLDEELNFLFNKEARNVTGDS--GPFLPRENGLGGGYQVLITPEYG 215
 :
Qy 420 YRDSLTLGPKNFSGIILDLLLELLELQELRHQLQWSPQAQRNI-----SLQDICY 472
 :
Db 216 --QDALAAP-----LLKHVEIMKHFIASFNVSHVGVDWSDSLSDICF 253
 :
Qy 473 APLNPD-----NTSLYD-----CC-INSLIQYFONNRTL-----LLL 503
 :
Db 254 KPAPPVAADSAASSLGDVIDKIVPCWIPTIDCFWEGSKALGHPHPSLPKSSLGPIGMILL 313
 :
Qy 504 TANQTLMGOTSQVDWKD-----HPLYCANAPLTF-----532
 :
Db 314 SS-----LSGDGMIRWSDFDPPIAVIDEIHRSFNLGSHYTPFERAGVSHGYMDRPCIDPLDP 369
 :
Qy 533 -----KDTALALSCHADYGAPVPFPFLAIGYKGKDYEAE 568
 :
Db 370 ECPPMAKNYFDVCPHIDRVREIAKKYCTELEEEKKKDGYSFFDFL---GRKKREAGDQP 426
 :
Qy 569 ALI-.I-----571
 :
Db 427 KMHPAPADSIPTIEDAPAQVPVSTAPIPTTTTSLSPPEARAAEEKKQKARELKDYC 486
 :
Qy 572 -----MTFSLNLYP-----AGDP 584
 :
Db 487 KSYRKSAPEMLKKKDKWPVMSNNYPQNVDYAAEMTGCGSFASNVLNWPEDMILGNP 546
 :
Qy 585 RLAQ--AKL-----WEAEFLER-MRA 602
 :
Db 547 RRACKGKGLSGDALQSVFLVSPADVFLRFKQPGKSNMKTGLDMAMNETAAEOVLQA 606
 :
Qy 603 FORRMAGMPQVTTFAERSL-----EDEINRTTAEDLPFATS-----YIV 642
 :
Db 607 WQR-----NFT--KSLYNHKANVEDGNERKTLH--PLASTSIADMLESFCQNYTI 654
 :
Qy 643 IF----LYISALGSYSMSRWVDSKATLG--LGGVAVVLGVAAMAAGPFSYLGISSSL 696
 :
Db 655 ILAGYALMLAYAIQTQARFDCNPATESMGLAGVLVVTTFASVAGLGATWFGIEFNA 714
 :
Qy 697 VILQVPLVLVSQADNIFTIVLEYQRLPRRPCEPREVHHGRALGRVAPSMLLCISEAI 756
 :
Db 715 ATTQIVPFLTILGIGVDNMFMLLHNRYRDVVVKLAGHAEMA-I--LMRETGMSILCTSINNII 772
 :
Qy 757 CFPLGALTMPAVRTEALTSLGAILDPLLQMSAFVALLSLSDKROEASRLDVCCCV--- 813
 :
Db 773 SFLTGTLLPIPALKRSFCAQSIIITFNFIATITYPATISIDLRRKKAQRDLVCCLYGD 832
 :
Qy 814 -----KPQELPPPQQGEGL-----LLGFFFOKAYA 837
 :
Db 833 TREESYMWISKPKIQSKRIICAGSEASIMQOPDGI TQAQWASSDDPAPWSLHSFIRYYI 892
 :
Qy 838 PFLHWTTRGVLLLFIALPFGSVLSYMCHTSVGLDQBELALPKDSYLLDYPFLNRYFEVG 897
 :

Db 159 KIEBELTVNSFPREAALEDNPALMKRKHLENETEIEENRNTI-----TFKD 205
Qy 485 CCINSLQYFQNNRTLLLTANQITMGQTSQVWDKHDPLYCNAPLTFKGTALALSCMA 544
Db 206 VCMWYGDYREKNVIELKRHELN-----KRGISVTPQVN 243
Qy 545 DYGARVPFPFLAIGG---YKGDYSEAEALIMTFSLNNYPAGDPRLAOKLWE---EAPLE 598
Db 244 QEGTPIYAFVIGGVDTPNDTIKYARANRLWYFLKPDDEQEQL---AKFWEDTAKEFVR 301
Qy 599 EMRAFORRMAGNFQVTFTAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGSYS--- 655
Db 302 ETYADHPFI---QCHIKHSRIVDQGLTRNANRLKPYENVVIAVLVLTAF---YSVKW 353
Qy 656 -----SWSRVWD---SKATLIGGGVAVLVGAVMAAGFFSYGISSVLIVQVPLV 706
Db 354 YPRDHSWP-LHIDWLRKPMALGSLVLAAILSGIGLLWFGMPFAEITL-IAPFLV 411
Qy 707 LSVGADNIFIFVLEY---QRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFLGA 762
Db 412 LSGVDDMFIAVAANHTEMTKYGRSPKWKGRMIEMSESAVALPITSFDVLSPGAGT 471
Qy 763 LTPMPAVTFTALTSGLAVILDFLLQWSAFVALLSLDSKQREASRLDVCCV----- 813
Db 472 ITDIAVQGFCAWTAACMFTEFLYQITTFEALMWISAKAQMGRNCPICITAGDIYIE 531
Qy 814 -----KPOELPP 820
Db 532 DGSLOPNLKKKSKRKOAKAEKKEKDEAKNDKNMEIENAEKSSYDSPDPQSQIHI 591
Qy 821 PQSGEGLLGFPQKAYAPLLHWTGTVGVVLLFLFGVSLVSMCHSHVGLDLOELALPKD 880
Db 592 PVKSRGAMGHFRDYVVPWMLLNWTKLFMFVTFIYILGISVIGICVMCGGLDYDKLLHS 651
Qy 881 SYLLDYFLFLNRYFVGAPVYFVTTLGYNFSSEAGNNAICSSAGCNFSFTQIKIYATEP 940
Db 652 DPLVE---ALKREIEL-----FHGGQJIEIA--- 674
Qy 941 PQSYLAIPASSWVDVDFDMLTFPSCCRLYISGPNKDKFCPTVNSLCKNCKMSITWGS 1000
Db 675 -----IQNCPNITIAE 685
Qy 1001 VRPSVEQPHKYLWFLNDRPNIKCPKG-----LAAVSTSVNLATSD-----GQ 1043
Db 686 SRDRIE-----LVAQEFENISYSGGKTSFWRVREYKYNLGSYLNDRNRESWIGV 738
Qy 1044 VLASRFMAYHK-----PLKNSQDY-TEALRAAR-----ELAAINTADLRKVP 1084
Db 739 YEWSQLFAFYKWLQSQDFWANESDYDTLELKSRYFRIGVHRLSTPTDLVL-ITEELRGVA 797
Qy 1085 GTDPAFEVPPYTTNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLS 1144
Db 798 DRHPDLNIVTYOOSRAIQANLVLSSTTNDTLAMPFCPCVALIFIPNPI-CALFITPA 856
Qy 1145 IVMILVDTCFMAWLDISYNAVSLINLVSAVGSVEFVSHITRSFAIS-TKPTWLERAKE 1203
Db 857 MVTIDIGVIGFLSWSKLDPIEMVITIMSIGSEIEFSAHITGFSVNDNSLSAPDRCDV 916
Qy 1204 ATISMSGSAVAGVAMTNLPGLVILGLAKAQLIQIPFRNLTLITLGLHLGLVFLPVILS 1263
Db 917 AMEKLAWPVHG-SLSITLGVFLAFIDSVMVLVFFKTSILVL-IIGAWHALMLLPIILS 974
Qy 1264 YVGP-----DVNPALALEQKRAEEAAVAVWVASC 1293
Db 975 MCPIVIERLSDASKASDRRLKLNKNKNSYVAINLP 1010

RESULT 15

S06119
membrane protein patched - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: S06119; A33468

R;Nakano, Y.; Guerrero, I.; Hidalgo, A.; Taylor, A.; Whittle, J.R.S.; Ingham, P.W.
Nature 341, 508-513, 1989
A;Title: A protein with several possible membrane-spanning domains encoded by the Drosop
A;Reference number: S06119; MUID:90015164; PMID:2797178
A;Accession: S06119
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-1299 <NAK>
A;Cross-references: UNIPROT:P18502; UNIPARC:UPI000014EA18; GB:X17558; NID:g8389; PIDN:CP
R;Hooper, J.E.; Scott, M.P.
Cell 59, 751-765, 1989
A;Title: The Drosophila patched gene encodes a putative membrane protein required for se
A;Reference number: A33468; MUID:90058658; PMID:2582494
A;Cross-references: UNIPARC:UPI0000165CA3; GB:M28418; GB:M28999; NID:g552097; PID:g55209
C;Genetics:
A;Gene: FlyBase:ptc
A;Cross-references: FlyBase:FBgn0003892
A;Map position: 2 4D3-D4
C;Superfamily: Drosophila membrane protein patched
C;Keywords: Glycoprotein; transmembrane protein
F;74-92/Domain: transmembrane #status predicted <TM01>
F;427-448/Domain: transmembrane #status predicted <TM02>
F;456-503/Domain: transmembrane #status predicted <TM03>
F;529-555/Domain: transmembrane #status predicted <TM04>
F;557-585/Domain: transmembrane #status predicted <TM05>
F;677-699/Domain: transmembrane #status predicted <TM06>
F;967-1017/Domain: transmembrane #status predicted <TM07>
F;1019-1047/Domain: transmembrane #status predicted <TM08>
F;1061-1086/Domain: transmembrane #status predicted <TM09>
F;1093-1121/Domain: transmembrane #status predicted <TM10>
F;142,238,335,388,807,861,1194,1271/Binding site: carbohydrate (asn) (covalent) #status

Query Match 6.7%; Score 464.5; DB 1; Length 1299;
Best Local Similarity 20.4%; Pred. No. 4.4e-24;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;

Qy 321 KTSLSDKLSFTHTLLGQFFQGWGTWVASWPLTILVLSVIVVALAAGLVFTLTTPV 380
Db 43 KKGARGRTAIIYLSVFSOSHLETGSSVQKAGKVLVAILVLTFCVGLKSAQIHSKVH 102
Qy 381 ELWSAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSRYOSLLGLKPNFSGIL 436
Db 103 QLMQEGGGLAEALAYTQKTIGEDBSATHLLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DLDDLLELLELQERLHLQVWSPEAQRNISLDICYAPLNPDNTSLY-----DCC 486
Db 154 HLEVIVKATAVKVHLYDTE-W-----GLRDMCNMESTPFEIGIYYIEQILRHLPDS 204
Qy 487 INSLIQYQN-----NRTLALLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCCWEGSQLGGSASVVIPLGNORLLTTLNPASVMQYKQKSEKISFDFTV 264
Qy 522 FLYCANAPLTFKDGFTALALSCW-----ADYGAPVFPFLAIG--GKY 560
Db 265 EQYMKRAAIA---SGYMEKPCINLPNCPDTPAPNKNSTQPPDVGA---ILSGGCGYA 317
Qy 561 GKDYSEAEALIMTFSLNNYPAGDPRLAQA----- 589
Db 318 AKHMWPEELIVGGAKRN-RSGHLRKAQALQSVOVQMTKEKMYDQWQNYKVHHIGWTQE 376
Qy 590 -----KLWEAF---LEEMRAFQRMAGMFOVTFTAERSLEDEINRTTAEDLPIFATSY 640
Db 377 KAAEVNNAQRNFSREVQQLLRKQRIATNYDIYVSSNAALDDILAKTSPSALSIVIGV 436
Qy 641 IVIFLYISIALGSYSWSRVWVDSKATLGLGVAVVLVGLAVMAAGFFSVLGRSSVLILQ 700
Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAAGLGLSALLGIVFVNAASTQ 492
Qy 701 VVPFVLVSGADNIFIVLEVYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFEL 760

```
Db 493 VPFLLALGVGHIFMLTAAYAESNR--EQTKL-----ILKKVGPSILFASCTAGSPPA 546
Qy 761 GALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
Db 547 AAFIPVPAKVFCQAIVMCSNLAAALVFPAMISLDRRTAGRADIPCCCFPWKEQ 606
Qy 815 PQELPP-----PQGGSGL-----LLGPF 832
Db 607 PKVAPPVLPLNNNGRGARHPKSCNNRVALPAQNPLLRADIPGSSHSLSASFSLATPA 666
Qy 833 QXAYAPFLHMTIRGVVLLLEALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLPLNR 892
Db 667 FQHYTFPLMRSWKPLTVNGFLAALISLSYASTRLQDGLDIIDLVPKOSNEHKFLDAQTR 726
Qy 893 YFEVGAVPVFTVTLTYNFSSEAGMNAICSSAGCNMPSFTQKIQYATEPPEQSYLAIP--- 949
Db 727 LF--GFYSYAVTQG-----NPEYPTQQLLEDY-HDSFVRVPHVI 764
Qy 950 -----ASSWDDDFIDMLTPSSCCRLYISGNPKOKFCPSTVNSLNCCLKC----- 993
Db 765 KNDNGGLPDFLLLFSEML-----GNLQKIFDEYRDGRLTKECWFPNASSDA 812
Qy 994 -----MSITWGSVRPSVEQ-----FHKYL-PWFLNDRNLIKCPKGG 1028
Db 813 ILAYKLIVGTGHVDNPVDKELVLTNRLVNSDGIINQAFYNLSAW-----ATNASSPTL 868
Qy 1029 LAA--YSTSVNLTSDQVLSARFMAHYHKPKNSQD-----YTEALRAAR 1070
Db 869 LRANCINREANGASQGLYPEPQYFHP--NEYDLKIPKSLPYVAQMPFYHLGTDTS 926
Qy 1071 ELAANITADLRKVGTDPAPFV--PPYTTINVFYQYLTLPGLFMLSCLVPTFAVSC 1128
Db 927 QI-KTLIGHIRDLSVKYEGFLPNYPGIPPFIFWEQYMT-LRSSLAMILACVLLAALVLV 984
Qy 1129 LILGLDLRSLNLLSIVMLVDTVGFMAIWDISYNAVSLINLVSAVMSVEFVSHITS 1188
Db 985 SULLLSVMAAVLVLSLASLAQIFGAMTLLGLKLSAIPAVLILSVGNMLCF--NLVLS 1042
Qy 1189 FAISTKPTWLERAKATISMSGSAVAGVAMTNLPGILVLGLAKAQLIQIFPRFLNLLITL 1248
Db 1043 LGFMTSVGNRRVQLSMQSLGPLVGHMLTSGVAVFMLSTSPFPFVIRHPCWLLLVLC 1102
Qy 1249 LGLHLGLVFLPVILSYGVDVNPALAEQRAEEAAVAVMVASCPNHPKSRVSTADNIYN 1308
Db 1103 VGACNSLLVPILLSWVGPEAE-LVPLE-----HPDRISTPSPLVR 1143
Qy 1309 HS 1310
Db 1144 SS 1145

RESULT 16
T21612
Hypothetical protein F31P6.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T21612
R:Percy, C.
submitted to the EMBL Data Library, March 1996
A:Reference number: Z19449
A:Accession: T21612
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-955 <WIL>
A:Cross-references: UNIPROT:Q19945; UNIPARC:UPI000007B318; EMBL:Z69884; PIDN:CAA93751.1;
A:Experimental source: clone F31P6
C:Genetics:
A:Gene: CESP:F31P6.5
A:Map position: X
A:Introns: 64/2; 141/3; 255/3; 308/3; 409/1; 435/3; 484/2; 579/1; 632/2; 715/3; 777/3; 8
Query Match 6.6%; Score 458.5; DB 2; Length 955;
```

```
Best Local Similarity 19.8%; Pred. No. 7.7e-24;
Matches 207; Conservative 207; Mismatches 402; Indels 227; Gaps 36;
Qy 336 LIGQFQFGWGTWASWPLTILVLSVIPVVALAAGLV-FTEL-TTDPVELWSAPNSQARSE 393
Db 9 LVYNAPFWGPVVRWRWFCFISPLFLTLACSVGFRTMELRVDDSYVFTTSDAARWR 68
Qy 394 KAFDQHGFGPFRTRNOVILTAPNRSYRYSLLGAPKNSFGILDLLELLEL-QLERLR 452
Db 69 ISVFNEW-PL--DENKFLPGKSPKAPFVNILIRAKDGGSIMRDNVLHIEILNQIMW 125
Qy 453 HLOWSPPEAQRNISIQDICYAPLNDNTSLYD--CCINSLLYQFQNNRTLLLTANOTLM 510
Db 126 NISITPDLKFLTYQDLCLS-----YDWVCGANEHIQMLRRNDV-----NQILD 171
Qy 511 GQTSQVDWMDKHPLYCAN--PLTFKQDGTALALSQWADYG-APVPPFLAIGGVKGDYSE 566
Db 172 LHPPRGCTKDTPVYLGIGFGDVQFQNGT-----LSDAKLTQLFYLKQDKQWVEYS 225
Qy 567 ABALIMTFSLNYPAGDPRLAQAQLWBRAPLEEMRAFQRRMAGMFQVTFABRSLEDTN 626
Db 226 K-----FSVAL-----ETFLNQVYS-----SDVITLSPAHYQSLDGLD 259
Qy 627 RTAEDLPIFATSYIVIFLYI---SLALGSYSWSRVWVDSKATLGLGVAVVLGAVMAA 683
Db 260 ENAKAFVFNFVVSFFVLAMALVSSFTLKSSATKIDMISSKPMWLAAGMFSTVLISIA 319
Qy 684 MGFFSYLGRSLVILQVVPPLVL-----SVGDNIPIFVLVYQRLPRRP 728
Db 320 FGFLILGVRYN-VINTIPIFLIIGEFQSFQMKENKQIPAIGDIDFLNACWDQTSKSL 378
Qy 729 GBPEVHIGRALGRVAPSMGLCSLEAICFFIGALTPMPAVRTFALTSLAVILDFLLQW 788
Db 379 SVPE--RASKTILSHAGVATTINVTDMVMSFAGCITDLPGIOFFCIYACVSAPSYQL 436
Qy 789 SAFVALLSL-----DSKQRE----- 803
Db 437 TPFSGAMAIMGEVEREKHCLFFYRTFQLVDISKNEEADSKLOQIKRSASAPFENYLS 496
Qy 804 -----ASRLDVCC--CVKPOLPPG-----QGEGLLGLPQKAYAPL 840
Db 497 SNSFSFSDSOSFSSKKTIPAEFAWKQOOSPNSLSKSKORKEKDRIVHFITGLYGPPI 556
Qy 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPV 900
Db 557 LNSVRIIFSGLIFFVYLAIAWYGCYNFREGLNPGNLVTDNDHYIAKYFSDIKHFWRIGAQL 616
Qy 901 YFVTTLGYNFSEAGMNAICSSAGCNMPSFTQKIQYATEPPEQSYLAIPASSWDDDFIDW 960
Db 617 HVAVLNPPLNLTISENRNEL-----LKVVSAPENTQY-TLREGTVTFPLEY 661
Qy 961 LTPSSCCRLYISGNPKOKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLPWFLNDRP 1020
Db 662 LN-----YLSELNAE-----VEDTERLWKTLKNSMDKY----- 689
Qy 1021 NIKPKGGLAAYSTSVNLT-SDGOVLASRFMAYHFKPLKNSQDYTBALRAARELAN---- 1075
Db 690 -----TGGSTQWASNLKINKTIDGSPQAFRQIALKNFVEPNDDHKAQLLRDIADHPEN 744
Qy 1076 -----ITADLRKVGTDPAPFVPPYTTNV-----FYEQYLTLP-----EGLFMLSCL 1120
Db 745 VVVYHEVSGNKKIL-NDPFISSHSCYAQKNIPKLAFFPADQYLIILPATIQNVVLSLLCM 803
Qy 1121 VPTFAVSCLLJGLDLRSLGALLNLLSVMLVDTVGFMAIWDISYNAVSLINLVSAVMSVE 1180
Db 804 -----AVVSFLLVPSLFGVFIIVSVISNIGVFGYMTLGVNLDVAVSMISIIMSGFAD 859
Qy 1181 FVSHITRSPALSTKPTWLERAKATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIPFP 1240
Db 860 LSAHIYAFVTSHGDT-KQRVIGALETLOWPFIQGSAST-INGISILVTVDAIIL-LVFP 916
Qy 1241 RINLITLGLLGLVFLPLVILS 1263
```

Db 917 KTIWLTMLGAIHGLFFIPIFELS 939

RESULT 17

A89153

protein C24B5.3 [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: A89153

R:anonymouse, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A:Reference number: A75000; MUID:99069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_elegans/

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999

A:Accession: A89153

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-956 <STO>

A:Crosses-references: UNIPROT:Q18129; UNIPARC:UPI0000076A25; GB:chr_V; PIDN:AAA96120.1; PIDN:CAA22312

C:Genetics:

A:Gene: C24B5.3

A:Map position: 5

Query Match 6.3%; Score 436; DB 2; Length 956;

Best Local Similarity 21.2%; Pred. No. 2.9e-22;

Matches 233; Conservative 177; Mismatches 353; Indels 338; Gaps 48;

QY 345 GTWASWPLTILVSVIPVVAALAG-LVFTELT--DPVELWSAPNSQARSEKAFDQH-- 400

Db 16 GENVHFRILYFLLSPAILTSAFAGFLWFTQTTNDPQVVFSPANAPWRYERAVLTHWP 75

QY 401 -----FGP---FFRTNQVLTAPNRSSVRYDSLLGPNFSGIOLDLLELLEQLER 452

Db 76 LDEKFWPKGSYDLHGVDVIAAGRIHPDY-----GRPN---ILNIRYLDVARINDYII 127

QY 453 HLOVSPQARNISL-----QDICYAPLNPDTSLVDCINSILQVFNQNRLLTLLT 504

Db 128 H-----NLTPVPDINGKHDIAYT-----DLCMRVDMACYLNDHITMLMP 167

QY 505 ANQ-----TLMGQTSQVDMKDFLYCAN 527

Db 168 KTRWGNFGPPFAELASDIINTQNTYPIGWRTETPIYFGLVAGPVLNVDSEGHFDYASA 227

QY 528 APLTP--KDGTLALSCMADYAGAPVFPFLAIGYKKG-----DYSEBALIMTSLNRY 579

Db 228 IRLTYNTRGK-----VDQYGT-----WRRKLSMWLTDKENPVSELEFGVNH- 271

QY 580 PAGDPLAQAKLWEAFLEEMRAFQRRMAGMFQVTFPFAERSLEDEINRTTAEGL-PIPAT 638

Db 272 -----NMTLPEGLQD-----VADTLAPKFGV 292

QY 639 SYTIVIFLY---ISLALGSSYSSMRVMVD---SKATILGLGGVAVVLGAVNMAAGFFSYLGI 692

Db 293 TCALITPFFSLVSVLRKHT--QVMPDVRSKPLVAAAGLWTPIMATVTSFGLIILWCGF 351

QY 693 RSSLVLQVFPVLVSVGADNIPFVLEYQRLPRRPGEPREVIHGRALGRVAPSMLLCSL 752

Db 352 LYN-AIVNVSFPLILCIGIDDLFIMCAEWHR--TNPQHSPEKRKIGKTLSEAAVAITSLS 408

QY 753 SEACIFELGALTMPAVRTFALTSGLAVILDLLQ-----MSAFVALL----- 795

Db 409 TDIATFAMGCYTLPVGMFCMTYCQCFCFYVIQIIFLGPVLAYAAHEQNGQHVLLLR 468

QY 796 -----SLDSKRQESRLDYCCCVKQELPP----- 820

Db 469 KAVDPDKTESPVKMLLSGSVNRQDQEARRN-----SRKQAVVEKEGKSKFGEVVE 522

QY 821 -----PGQG-----EGLLIGFTQKAYAPFLLHWITRGVVLLFLALFGVLSYMSCH 866

Db 523 KLEHTLEKHHDDPGHNSSEETLVSKVFRITIGPFILOKSTQVCALLLYLVYISIAOGCLN 582

QY 867 ISVGLQDELALPKDSVLLDYFLFNLN-YPEVGAPVYFV-----TTLYGNFSSEA 914

Db 583 IKEGLDPKLLVRESFYLSKFVEIIDETFWREGLQVQVWVNNPPDLFTPTETRKGF----- 637

QY 915 GMAICSSAGCNPFSTQKIYATEFPEQSVLAIPASSWVDDFDLWLTSPSCCRLYISGP 974

Db 638 -----EMMAEPENTQYTHPNATMI-----WL-----RAY----- 662

QY 975 NKDKFCPSTVNSLNCNCKNSITWGSVRPSVEQPHKYLPLWFLNDRPNIKCPKGGGLAAYST 1034

Db 663 --EHLLETEVHELNIK-----PNSVTVEWTRCDWLI-----VAGGRLLWQM 703

QY 1035 SV---NLTSQ--GOVLASRFMAYHKLKNSQDYTEALRAARELAANITADLRKVEGTDPA 1089

Db 704 DMVWANTTETPRITAFRQGLRNYRTPTDHTHSCKLMBSIA-----DK 749

QY 1090 FEVPFYITNVFY---BOYLTLPEGLP--MLS-LCLVPTFAVSCLLGLDLRSGLLNLL 1143

Db 750 YSMFNVTTFHEYYPFADQYLELP-SLFQNMISDLCTI--FAVSWVMIP-ELICAVAVL 805

QY 1144 SIVMILVDTVCFMAMWDISYNNAVSLINLVASVMSVEFVSHITRSFALSTKPTWLERAKE 1203

Db 806 SIASINVGVLGFMGFWGNLDSVSIITVICIGFSVDLSAHIAVAFSQSYGNSH-ARAVA 864

QY 1204 ATISGSAVFAGVAMTNLPGILVLGLAKAQLIQIPFPRNLNLLITLGLLHGLVFLPVILS 1263

Db 865 ALETLGMPVFLGASSTVL-GILLTLVDSYIVQFFKTVFLVIN-FSILHGLIFLPIFLM 922

QY 1264 YVGPDPNPALALEOKRAEEAV 1284

Db 923 KVVREVK-----ESPRDDQSV 939

RESULT 18

T26521

hypothetical protein Y18D10A.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T26521

R:Haris, B.

submitted to the EMBL Data Library, December 1998

A:Reference number: Z20226

A:Accession: T26521

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-936 <WII>

A:Crosses-references: UNIPROT:Q9XW22; UNIPARC:UPI00000813DB; EMBL:AL034393; PIDN:CAA22312

A:Experimental source: clone Y18D10A

C:Genetics:

A:Gene: CESP:Y18D10A.7

A:Introns: 11/1, 44/2, 106/1, 166/3, 289/2, 373/1, 396/3, 485/1, 573/3, 615/3, 644/3, 65

Query Match 5.9%; Score 409.5; DB 2; Length 936;

Best Local Similarity 20.2%; Pred. No. 2e-20;

Matches 213; Conservative 172; Mismatches 369; Indels 301; Gaps 40;

QY 342 CQSGT-----WVASWPLTILVSVIPVVAALAGVFTLTDPVELWSAP-NSQARSE 393

Db 29 QPWANIVAKYCLFVAKYWPFFIILITICLSMGIILNFKIVRGVNYLYAPLNATWKE 88

QY 394 KAF-----HDQHPGPPFRTNQVILTAPNRSSYVDS--LLLGPKNFSGIILDLLEL 444

Db 89 EAVFGENWAKODDF-----YPGKDIRRQGIYLVNNAKGGNVLNRQEAQDF 136

QY 445 LEQLERHLQVWSEAPQNRNISLQDICVAPLNPDTSLVDCINSILQVFNQNRLLTLLT 504

Db 137 LKILDMILNKLSS-SAGRIFTYKQVDC---LHFQN-----DC-----FSPHAKLL-- 178

QY 505 ANQTLMGQTSQVDMKDFLYCANAPLTFKDGTLALSCMADYCAPVFPFLAIGYKGDY 564

Db 179 -----ANIYSKHQNSMENITVPIRST-----YATEPIDISKVLGNVSLDY 220

QY 565 -----SEABALIMTFSLNN-----YPAGDPLRAQAKLW---BEAFLEEMR 601

Db	221	DGHVENASAWLILYQLKNEKQWLSRDFEDGLAEKIQSGEAPSELLNLYYFHSATFDQBLE	280
Qy	602	AFQRMAGMQVFTTFAERSLEDEINRTTAEDLPFA--TSYIVIFLYISALG-----SYS	655
Db	281	KENRRLTTPKFSITFSV-----LIIFAITMTFTTKPKMKTENGINQYPI	325
Qy	656	SHSRVMVDSKATILGLGAVVLGAVMAAGPFSYLGIRSSLVLQVPPFLVLSVGDNIF	715
Db	326	DM-----VLSKPLIGCGVLTWCATIISSGLLMLFNV--TFVDMCTVMPFLSTIGIDDTF	380
Qy	716	IFVLEYORLPFRPGEPREVHIGRALGRVAPSMLLCSLSAICFFFLGATPMPAVRFALT	775
Db	381	LMLAAWHEYDR--NLPEYKRIEKAMRHAASVITSITLDALAFGLIGSIAPLPAVIFYCY	438
Qy	776	SGLAVIDLFLQMSAFVALLSLSKQASRLDVC-----	810
Db	439	SSAAILFIFYLVITMFAVLAQGREEDLKHSVTGMKTIDLDSDYETASTRQLLLKMS	498
Qy	811	CCVKPQELPPQGGGLLIG-----FPQKAYAPFLHWHITRGVLLFLALP	857
Db	499	RVSVKADEENNNNNEKSIENTIKIDNRMYORFFEDQYAFISNSKISILSFLIYLAYL	558
Qy	858	GVSLSMCHISVGLD-----QSLALPKDSYLLD	885
Db	559	AAAFYGVKRLKIGDFPVTVVTLIKYKVLQINIVQEDSASRVFLEVRQLPEDTKLMD	618
Qy	886	YFLFLNRYFEGVAPYFVTTLYGYNFSSEAGMNAICSSAGCNPFSFTQKIYATEPPEQSY	945
Db	619	IAV-----MNSP-----NFSNPE-----ERNP-----NEVLSEP-----	643
Qy	946	LAIPASSWDDFDID-----WLTSSCCRLYISGNKDKFCPSTVNSLNCUKNCMSITMGSV	1001
Db	644	----ESTWCSEGRESTQFWFFE--MQYLSNLGFGGDLTKMNS-----ERKLSQSKTF	692
Qy	1002	RPSVEQPHKYPFLNDRPNIKCPKGLAAYSTSNLTSQGV--LASRPMAYHKLPKNSQ	1060
Db	693	LMSHEKF-----GVDVSDKQFRLSTR-----LKNVE	719
Qy	1061	DYTEALRAARELAANITADLRKVPDPAFEPYPTITNVFQYVLTILPEGLFMLSCL	1120
Db	720	TDEMFENCAR-----TWKLSQKHANYSIITVSLPNIADIEDIMFPQTMQDIYSI	771
Qy	1121	VPTFAVSCLLGLDLRSLGILLNLSIVMILVDTVGFMAWDISYNAVSLINLVSAGMSVE	1180
Db	772	AVMVPVALLFIPOPLCSVIIG--LNTASIAFGVIGTMSFLGSLDATSMITVAMSVGFSVD	830
Qy	1181	FVSHITRSFAITKPTWLERAKENTISMGSAPFA-----GVAMTNLPGLVLGLAK	1231
Db	831	FAAHV--SVAYMTE-----SGAQIPGKSATYSRPFCHTLGTIGMPVTOASVSVLLGVSS	881
Qy	1232	AOLIQIFF-----FRNLNLTLLGLLHGLVFLPVIL	1262
Db	882	LVLDSVYVQTCFRTVVLVILFGTTHALVFLPILL	916
RESULT 19			
T26683			
hypothetical protein Y38F1A.3 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004			
C:Accession: T26683			
submitted to the EMBL Data Library, October 1998			
A:Reference number: Z20253			
A:Accession: T26683			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-800 <WIL>			
A:Cross-references: UNIPROT:O3XWL9; UNIPARC:UPI000007721A; EMBL:AL032639; PIDN:CAA21636			
A:Experimental source: Clone Y38F1A			
C:Genetics:			
A:Gene: CESP:Y38F1A.3			
A:Map position: 2			

A;Introns: 13/2; 100/2; 132/3; 220/3; 257/2; 521/2; 642/3; 710/3			
Query Match 5.9%; Score 407; DB 2; Length 800;			
Best Local Similarity 23.4%; Pred. No. 2.5e-20;			
Matches 175; Conservative 119; Mismatches 285; Indels 168; Gaps 27;			
Qy	594	EAFLEMRAPQRMAG-----MPQVTFTAERSLEDEINRTTAEDLPFAFATSIVIVP	644
Db	121	ETMYIMKEWEOKLPAYTLSTENDPLIRVYVTTSEGLVSEVRRTGILAMPLMGVTEPLII-	179
Qy	645	LYISALGSYSWSMVDSKATILGLGVAVVLGAVMAAGPFSYLGIRSSVLVLOVVP	704
Db	180	----LAFITLITLTKDDPVKSKPFEAFIAGVICILSCASFGHLFWMGF-EYIPVITVVP	234
Qy	705	LVLVSAGDNIFIFVLEYORLPFRPGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALT	764
Db	235	LILSIGVDDVFIHAWHRTPYK--HSVRDRMAETLADAGPSISITSLNLSFAIGIT	292
Qy	765	PMPAVRTTALTSLGLAVILDPLQMSAFVALLSLSKQASRLDVCCKVKPOELP--PPG	822
Db	293	PTPAIYTFCVFISTAVIYDIYQIFPFSAVLVLSGEREE--QRKNAYLMMKDVPLPEPTG	351
Qy	823	QGE-----GLLIG-----PFQKAYAPFL-----HWI--TRGVVLLFLALFGVSL	861
Db	352	KOKKETIVSRVLGKVDLFWVDFIMAKWSKFLIGAIMITYIMFARGVM-----	399
Qy	862	YSMCHISVGLDQELALPKDSYLLDYF--LFLNRYFEGVAPV-YFVTTLYGYNFSSEAGMNAI	919
Db	400	----EIAVGLSEKFLDSDSPLLPLVRQTNTVIFKEGGQVAVFVNNPG--NMSEPDVAPPI	454
Qy	920	CSSAGCNPFSFTQKIYATEPPEQSYLAIPASSWVDDFDIMLTSPSCCRLYISGNPKKF	979
Db	455	-----MRILRRFEVANNVSGAASTHMLLPYLPYVGEQE-----	488
Qy	980	CPSTVNSLNCUKNCMSITMGSVRPSVEQPHKYPFLNDRPNIKCPKGLAAYSTSNLT	1039
Db	489	-----HGSIEFKYRIPLPEFFK-----LTFRRWSHFVNGL	518
Qy	1040	SDGQVLASR-----FMAYHKPLKNSQDYTEALRAARELAA-----NITADLRKVP	1085
Db	519	NHQDCLSEKPSCLQKFVSTGFHDVSW--DRLLLENWREMASEYQHLNLT-----	570
Qy	1086	TDPAFEVFPYTTINVFYQYLTILP---EGLFMLSCLVPTTFAVSCLLGLDLRSLGLNL	1142
Db	571	-----YEDFSMYSQDLTIIVPTQSTVFCALICMI-----MILTPTSPVTV	614
Qy	1143	LSIVMIL---VDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSFAITKPTWLE	1199
Db	615	TSTAVALSINLGVFCGLVYMNIDLDPIGNTTLLMAIGFSVDFAHITWHYYKGFHSHKRA	674
Qy	1200	RAKEATISMGSAPVAGVAMTNLPGLVLGLAKAQIQLIPFFPLNLTLLGLLHGLVFLP	1259
Db	675	RIRHALAGIAMPFQAGTSTML-AITVLALVHAYMVQV--FVKVVLVIFLGMFHLGLVLP	732
Qy	1260	VILSYV-----GPDVNPALAEQKRA	1280
Db	733	IVFSALPFTKTSQPQKKVAPLQMEHA	759
RESULT 20			
T31739			
hypothetical protein T21H3.2 - Caenorhabditis elegans			
C:Species: Caenorhabditis elegans			
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004			
C:Accession: T31739			
submitted to the EMBL Data Library, July 1997			
A:Description: The sequence of C. elegans cosmid T21H3.			
A:Reference number: Z21077			
A:Accession: T31739			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-881 <SCH>			

Db 628 TFAFKISNNIRTSLLLTWRNITSSYPEDDALVFDENNYSQDMLESLQSTILSSLTGTAII 687
QY 1115 MLS-LCLV-----PTFAVSCILGLGLDLSGLNLNLSIVMLVDTVGFMAWDISYNAYS 1167
Db 688 LLSAICIMIAESYIVFVVVCTILSND-----IGTAGILSLWGADLDPM 732
QY 1168 LINLVSAGMSVEFVSHITRSPFAISKPTWLRKAETISMG-SAVFAGVAMTNLPGILV 1226
Db 733 VVNILMSIGQCIDFATHGIRIYRSEYSDPDERIDAMGAIGWPFVVOAGT--STLLAIVV 790
QY 1227 LGLAKAQLQIPFERLNLATLGLLHGLVFLPVIL 1262
Db 791 MLMVPSSARM-FARTSVLVVGTGPFPHGLIIPV 825

RESULT 22
T16126
hypothetical protein F21H12.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T16126
R:Pavello, T.
A:Description: The sequence of C. elegans cosmid F21H12.
A:Reference number: Z18464
A:Accession: T16126
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-714 <PAV>
A:Cross-references: UNIPROT:Q09540; UNIPARC:UPI0000076F08; EMBL:U23176; NID:G726404; PID
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F21H12.4
A:Introns: 25/1; 334/2; 430/2; 543/3

Query Match 5.5%; Score 379; DB 2; Length 714;
Best Local Similarity 23.0%; Pred. No. 2e-18;
Matches 173; Conservative 119; Mismatches 299; Indels 160; Gaps 25;

QY 545 DYGAPEPPLATGGYKDYSEBALIMTFSLN--NYP-----AGDPLAQ--AKLWERAP 596
Db 58 DYAEW-----TGCGSG-----FASVLNWPEDMILGNPRRAKGGKLSGADA 100
QY 597 LEEMAFQRRMAGMFOVTFABRSLEDEINRTTAEDELPIFATSYIVIFLY-----646
Db 101 LQ-----FVPLVASPADVFLRFKQSLER-----TLETLSLYEHCHLHRHPLLCRSNSV 151
QY 647 -----ISLALGSYSWSRVNV-----DSKATILGGVAVVLGAVMAAMGFFSYL 690
Db 152 NSTPSSLLDMLSCWHMLLLILDLTAPRLQNHQWGFALGVVTFASVAGIGLATWF 211
QY 691 GIRSSVILQVPELVLSGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVPSMLLC 750
Db 212 GLEFNAATQIVFFTLGTIGVDNMFMNLNRYDVVKLAGHAEMAI--LMRETGMSILCT 269
QY 751 SLSEAICFPLGALTMPAVRTFALTSGLAVIDLLOMSAFVALLSLSKROEASRLDVC 810
Db 270 SINNLISFLTGTLLPILALRSFCAQSSILLTNFTAILTIYPAIISIDLRKKAQRDDP 329
QY 811 CCVK-PQELPPPPQEGEL-----LLGFFQKAYAPFLHWTIRGVVLLLP 854
Db 330 TREBSTSEASIMQFDGITOQAQWASSDDDPAPWSLSHFIRYIYIPFISKPAKVAIIVGCC 389
QY 855 ALPGVLSYMSCHISVGLDDELALPKDSYLLDYFLNRYFEGAPVYFTTIGYNS----911
Db 390 ALLGASFVGMQROSTLGLGLDVLPEHTAPAQFLRADKDYFSP-YPM-FAVIRKPNIDYAH 447
QY 912 SEAGMAICSSAGCNFSPQKIQYATEPEQSYLAIIPASSWVDVDFIDMLTSPSSCRLYI 971
Db 448 QQRQIDNTYKQISGSSKIYVTKN-----NEEPSEKY-----WGLMRDLI--STORGPD 494
QY 972 SGNPKDKF-----CPSTVNSLNLKNCWSITMGSVRPSVQFHK--YLPWFLNDRPNI 1022

Db 495 BEVAGSGFNLTSGTWTIGSVNSVEDARLAHALMCSHGLSFPACAGRIGKIRLIPFYL-----549
QY 1023 KCPKGLAAYSVSNLTSDQVLAERFMAYHKPLKNSQDYTEALRAARELANITADLRK 1082
Db 550 -----GLTDTAVIVDAIKDIRSVCRP-----571
QY 1083 VFGTDPAPEFVPTTNTNPFYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLSGLLNL 1142
Db 572 ---TQGLPNFPFOGIAFTFWEQYFLTGNLMQAISITITISVFCVISVLL-FNPWAAALMVV 627
QY 1143 LSIIVMLVDVTCFPMALWISYNAVSLINLVSAGMSVEFVSHITRSPFAISKPTWLRKA 1202
Db 628 CYLGIMTCELAGFMGLVGIKLPVSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
QY 1203 EATISMGSAVPAGV---AMTNLPGLILVLGLA 1230
Db 683 QRTSSAVDRVFVPIHGSFSTLLGILMLGFS 713

RESULT 23
T23399
hypothetical protein K07C10.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T23399
R:Coles, L.
A:submitted to the EMBL Data Library, December 1994
A:Reference number: Z19736
A:Accession: T23399
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-690 <WIL>
A:Cross-references: UNIPROT:Q09934; UNIPARC:UPI00000835CB; EMBL:Z47074; PIDN:CAA87375.1
A:Experimental source: clone K07C10
C:Genetics:
A:Gene: CESP:K07C10.1
A:Map position: 2
A:Introns: 34/3; 362/3; 448/2; 507/3; 548/1; 614/3

Query Match 5.3%; Score 367; DB 2; Length 690;
Best Local Similarity 21.3%; Pred. No. 1.3e-17;
Matches 163; Conservative 148; Mismatches 314; Indels 140; Gaps 27;

QY 564 YSEABALIMTFSLNNV-----PAGDPLAQ-AKLW-----BEAFLEEMRAFORMA 608
Db 8 YPLAARFGHPFYLGNQPFQVNVNGQASKAGPIKTAKEFVALWYMSKAETP-EQKQKQAVQL 66
QY 609 GMFOVTTAERSLE-----DEINRTTAEDELPIFATSYIVIFLYISLALGSYS 656
Db 67 GIFKKSARPNNLQPDFEMFGDQVANSEMLRGTLTTVKLFFIGGCLMWAFMACTFTELTI 126
QY 657 WSRVWVDSKATILGLGVAVVLGAVMAAMGFFSYLGIIRSLVILQVVPFLVLSVGADNIFI 716
Db 127 FSKOMLIAGA---IGSPLAATGACPAIILGWVGH-----PFNSIMCITPFLILGIGVDDAFL 179
QY 717 FVLEYQRL--PRRPGEPREVIHGRALGRVAPSMALLCSLSEAICFFLGALTMPAVRTFAL 774
Db 180 LINCNRREGKDKSAKQAEQLARVIREISPSMAITSLTNTNWFVGFLAPTPQMSFCL 239
QY 775 TSGLAVIDLFLQMSAFVALLSLSKROE---ASRLDYCCCVKPOELPPPGQEGILLGF 831
Db 240 GTALAIVLDFLEFLIFVPCWVLFYKKKEPIIAIENEERFANEKTEKRSSTSTGKYSRINW 299
QY 832 FQKAYAPFLHWTIRGVVLLFLALFGVSLYSMCHISVGLDDELALPKDSYLLDYFLN 891
Db 300 --KCFNLLLSVPGRVLVILIYFSIFTASYFGVAKMEKTFDPSKTFPPSDSKLVDLSASF 357
QY 892 RYFEGAPVYFTTIG--YNFSSEAGMAICSSAGCNFSPQKIQYATEPEQSYLAI 949
Db 358 SIQVYSPLNFLSSVFDLKNKDTDVANFN-----EMLHKLFPREECYGDVG 402
QY 950 ASSWVDDDFIDMLTSPSSCRLYISGPNKDKFCPSTVNSL-NCLKNCWISITMGSVRPSVEQF 1008

QY 620 SLEDEINRTAEDLPPIFATSYIVIFLYISIALGSSYSSRWVMDSKATILGLGVAVVLGA 679
 Db 188 VLDNEMKQGRRLTPFAAGFGEITFVSLCVILTRAVYINCLDQGGKLLISLGAVALCPILA 247
 QY 680 VMAAMGFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGRA 739
 Db 248 IYSTYGVILGNRTNSFML-VMPFLVMGIGVDSCLFMTHSWOKERRAQAASGTGNRLGMV 306
 QY 740 LGRVAPSMGLCSSEACFFLGLATPMPAVRTPALTSGLAVILDFLLQMSAFVALLSLDS 799
 Db 307 YESVGPSITITSLTDFLSFAIGALAPTPTBLPCIASIALALTYILQLVLFQPIALAVAT 366
 QY 800 KQCEASRLDVC---CVKPEQLPPQGEGLLGFQ-----KAYAPFLHWTITRGV 848
 Db 367 RYEHKTTITTSNCRWKRVRIR-----LGDGIIFSKLKAKCKKCNVYCKILSNKFPFAV 420
 QY 849 VLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYEPCVAPVVFVTLIGY 908
 Db 421 VMLGTAIYVFAIYGLMTWKTRLDVAKIILPKDSPLQRPNLVLTNL--VWANTHPVTIL-- 476
 QY 909 NPSSEAGMNAICSSAGCNFSTQKIYATEPPEQSYLEAPAS--SWVDDFDIMLTPSSC 966
 Db 477 -----INAPLDLE--NRHOMDRYWNWVDEPEKMHCKGKASTLSLWLRDIKFSYHGEP 527
 QY 967 CRLYISGPNKDFCEPSTVNSLNCMSITGWSVRPSV-BOFHKY-----LPWFLN 1017
 Db 528 FNLF-----AFFGLMTPEVYEEVDYDQANITTAFLKPEFLK 562
 QY 1018 DRPNIKCPKGLAAYSTVNL-TSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANI 1076
 Db 563 S-PFPK-----HMSDFHYHVEGVIKLDRFM-----MNWAYONT-----SSWDTRIQL 605
 QY 1077 TADLRKVPCTPAFEPFPTIITNVFYEQL-----TILPEGLFMLSCLVPTFAVSCLLIG 1132
 Db 606 MTDWRKVANNYSDLNVVWEPNGMFDQMLSLGRATATQGIWTL-VCM---AVVCAIFI 660
 QY 1133 LDIRSGLNLLSIVMLVDTGFMALWDISYNAVSLINLVSAVMSVFPVSHITRSFALS 1192
 Db 661 PNPCSIITATVSIASITITGVMGFLSLWSPDLDPVWMAAVLMSIGLSLVDPIAHVAYHFQLA 720
 QY 1193 TKPTMLERAKEATISM-GSA-----VPAGVAMTNLPGIILVGLAKAQILQI 1237
 Db 721 HRKE-IRNGKIKKIPKUGSTERLEHTLGAVAMPQAGV--STICILPL-LFRASYSPS 776
 QY 1238 FFRLMLLITLGLHLGVFLPVILSYVGPVNPA 1272
 Db 777 VFVAFLVVTGMLHGLLILPTFLAALPESVTTA 811
 RESULT 26
 T32908
 hypothetical protein F56C11.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T32908
 R:Tin-Wollam, A.; Wohldmann, P.; Morris, M.
 A:Submitted to the EMBL Data Library, January 1998
 A:Description: The sequence of C. elegans cosmid F56C11.
 A:Reference number: Z21244
 A:Accession: T32908
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-820 <TIN>
 A:Cross-references: UNIPROT:044978; UNIPARC:UPI000007C788; EMBL:AF043697; PIDN:AAB97556.
 A:Experimental source: strain Bristol N2; clone F56C11
 C:Genetics:
 A:Gene: CESP:F56C11.2
 A:Map position: 1
 A:Introns: 24/1; 58/2; 97/3; 134/3; 171/1; 200/3; 332/2; 368/3; 479/2; 503/3; 580/2; 678/2

Query Match

Best Local Similarity 5.0%; Score 347.5; DB 2; Length 820;

Matches 198; Conservative 161; Mismatches 390; Indels 219; Gaps 41;

RESULT 27

T22186

hypothetical protein F44F4.4 - Caenorhabditis elegans

QY 343 GW-----GTWVAGWPLTILVLSVIPVVALAAGLVFTLTTPVLMWSAPNSQARSEKA 395
 Db 2 GMDVISAVVARLAAVPVYSILLSILVLSGLVHRLBPDIRKAFSPDDSDAGVETR 61
 QY 396 FHDQHFPGPPFRNOVILTAPNRSSVRYDSLLGPKNPFSGILDLDLLELLELQERHLQ 455
 Db 62 VWLEYG-----LDIYPERAF-----CIPTAKRENSILOEGALADITYVDNRL----- 105
 QY 456 VMSPEAQRNISLQDICYAPLNPDNTSLYD--CCINSLQYFQNNRLLTLLLTANQTLWGOT 513
 Db 106 -----SAAVGQD-----GDGRKNCDPCLDLSPPHLLAN-----LTRNN--GTT 144
 QY 514 SQVDWKDHFLYCANAPLTPKD--GTALALSCMADYGAPEFPFLAIGYGKGYKDYSEABALIM 572
 Db 145 SVFTYPD-----MPYSGLDIFLGLHLS--NADFPPT-----ERISSKSLVLM 185
 QY 573 TFSLNYPAGDPLRLAQAKLWEAFLEEMKRAPORRMAGMFQVTF--AESLSDEINRTTA 630
 Db 186 YFSRSDTPEG-----KLAFKDAIDELQFSKSSAFEHVQVTFIPSDQVANREMRIGAI 238
 QY 631 EDLPFATSYIVIFLYISIALGSSYSSRWVMDSKATILGLGVAVVVLGAVMAAMGPFPSYL 690
 Db 239 EATLTMTIGFFLLLTQVILIVRLSS-----VGMAYLVATSLTPTMAATVASFCAICWM 293
 QY 691 GIRSSLVILQVVPFLVLSVGADNIFIFVLEYQ-----LPRRPGEPREHVHIGRALGRVA 744
 Db 294 GF-PSFSIQCVTPFLVIGIGVDDAFILHRWKHHTAITDTPR-----LEQVIVDVG 344
 QY 745 PSMGLCSSEACFFLGLATPMPAVRTPALTSGLAVILDFLLQMSAFVALLSL----- 797
 Db 345 PSITITSLNITIAFGIGFLTPTPQMSLFLCTASLALLLDYITFTYITILABIVELCNDPDYK 404
 QY 798 ---DSKROEASRLDVCCKVKPELPPQGEGLLGFQKAYAPFLHWTITGVVLLFL 854
 Db 405 STANTEPEKISRADT-----WLSKYSFSCVSKVRACIMVLL 442
 QY 855 ALFGVLSYMSCHISVGLDQELALPKDSYLLDYFL-----PLNRYEFGAPVYFVTLIGYNP 910
 Db 443 GMYALTITGVVMTKTTPPAKAFPSNKLVDLSNIPKVFNYF---PIITVINPPNI 498
 QY 911 SSEAGMNAICSSAGCNFSTQKIYATEPPEQSYLEAPASWVDDFDIMLTPSSCCRLY 970
 Db 499 EIKAEYDSF-----NN--MMDRLEH-----VPG----- 519
 QY 971 ISGPNKDK-FCPSTVNSLNCMSITMG-SVRPSVQFHKYLP-WF--LNDRNKICP 1025
 Db 520 IRGDNRSILIFQYVN-FDKTNILSTLVGEKYPISYN-----LPTMDAIGNPPLVVKYH 574
 QY 1026 KGGLAAYSTVNLTSDGQVLASRFMAVHKPL--KNSQDYTEALRAARELAANITADLRKV 1083
 Db 575 MG-----DDNKTIVT--APRLTLGKMSWAERARAKQH-----IRTI 611
 QY 1084 PGTDPAFVFPVPTIITNVFYEQLTILPEGLFMLSCLVPTFAVSCLLIGLDRSGLNLL 1143
 Db 612 LHEESQFNATLPDCDSAILSILVTGTLIGSIATVTV-CMAIVCFVFTANFNAVAVITS 670
 QY 1144 STVMLVDTGFMALWDISYNAVSLINLVSAVMSVFPVSHITRSFATSKPTMLERAKE 1203
 Db 671 VIASICYVVLVGLSLWGADLPVIOVDVLLATGFSVDYTAHVAYNF-FRARTPTQERVYS 729
 QY 1204 ATISMGSAVF-AGVA--MTNLPGLVVLGLAKAQILQIPFRLNLLITLGLHLGVFLPV 1260
 Db 730 SLAEMAMPCEAGLSTFLCMLPLIFVPTYA-----IVCFAKTVFLVAIGLHGLFILPV 784
 QY 1261 ILSYVGPD 1268
 Db 785 ILALPSKD 792

Query Match	2.7%; Score 189.5; DB 2; Length 1154;
Best Local Similarity	22.0%; Pred. No. 7e-05;
Matches	137; Conservative 87; Mismatches 212; Indels 187; Gaps 28;
Qy	341 FQHGCTWVASWPLTILVLSPVVALAAGLVFTELITD-----PVELWSAPNSQ 389
Db	23 PSRYGTWTAHRKVTGLPSLVALVFLCLYPIFFL-YTTDATNTVTSVGNLPHHWMT--DAQ 79
Qy	390 ARSEKAFHDQHFGFPFRTNQVILTAPNKRSSVRYDLSLLGPKNFGSGLLDLLELLELQ 449
Db	80 PLSERAVAEP-----DVIM-----RSVWVHGS-----YKALDRILGALQLQD 119
Qy	450 RL-----RHLQWSPEAQRNISLODICYAPLNPNQNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db	120 TLLGPTTNFNPRRPHASKIL-----PHDP-----TVDMTRDQ-----DA 154
Qy	506 NQTLMGQTSQVDWKDFLYCANAPLTFQDGTALALSCADYGAPV-----550
Db	155 FHVINGLTQ-SWFFH-----SPLOWGGSDAINIADSDIISTVNEKKTQSTSVNITLR 207
Qy	551 FPFLAIGYKGYKOYSE-----AEALIMT-FSLANNYPAGDPRLAQAKLWEBAFLEEMRAFQ 604
Db	208 HSPV-----FSGKRFEEURLVAADALVITLHLKXDSPLGRQWVNAE-----249
Qy	605 RRMAGMFPQVFTT-----AERSLDEINRRTAEDLPIFATSYIVIFLYISLALGSYS 655
Db	250 -RLAAQVHDKWTVIPHDGKSWASOLYEFQFPMSWPDWFLTTAYSITLTFYLLLSKLR 308
Qy	656 SWSRVMVDSKATLGLGGVAVVLGAVMAAMG-----FFSYLGIRSLVILQVVPFLVSVGA 711
Db	309 A-----VKSRFGL--MATILAQAASIGSSFTICAIFKIDLSRPVYAYVPLVWLAIMS 359

Db 175 MLPEDARNGLVSRDYS---STIIIVSNRREKQALVRVYNEIKNEIEKNFPPEG----- 226

Qy 588 OAKLWEEAFLEEMRAFORMAGMFQVTFTAERSLEDEINRTTAEDLPIDFATSYIVFLXI 647

Db 227 -----VEVVLTGDLGIITYKILEMDNDNRMTA-----VAGIIVILLILYP 267

Qy 648 SLALGSYSSRSRVWDSKATGLGGVAVLGVANMAAGPFSYLIGIRSLVLQVPPLVL 707

Db 268 -----YKSPIRMELVPL-----IPLIFGVTV-TLGFMLGLIPLD-IATSVTGAMII 311

Qy 708 SVGAD-NIFIPLVLYORLPRRPBPREVHIGRALGRVAPSMLLCSLSSEAICFFLGALTPM 766

Db 312 GKGIDYGHVHTNRYER--RNKGRSPBEABEASETGKALLGAALTITIAGFIALSLSIL 369

Qy 767 PAVR--TFALTSGLAVIDLFLOMSAFVALISLDSKRQEASRLDVCCVFQELPPPGQG 824

Db 370 PSLKRLSLSLMVGLAAVNNAVITPALAIL-----EEEFRTKI--LGKKGEIISIGE 421

Qy 825 EGLL-----LGPFQKAYAPFLHLHWIRGVVLLFLALFGVSLSYMCHHSVGLDQDELALP 878

Db 422 EGKLGFIFSRLGLFIKR-SPWL-----ALLIADVIMISGVSLYGATKITTEVRLEKMIP 472

Qy 879 KDSVLL-----DYFLNRYPEVGAPVYFTTLGY--NFSSBAGNAICSSA 923

Db 473 MDMPHQALTDIRSEFGGQDEITILIKADDVRDPETVRSILRFEREVKADSYNNVFESE 532

Qy 924 GCNNFPTQKIQYATEPPEQSYLEIPASSWWDDFDIDLTPSSCCRLYISGNKDCKPCPT 983

Db 533 SIADV-VIEKYGYTPEDKEIAKALEGSSLV-----SS 564

Qy 984 VNSLNCUK----NCMSITMGVSRRPSVEQFKHYLPFWLNDLRPNIKCPKGGLAAYSTSVMNLS 1040

Db 565 DYSLLTIKVLKGNFMGVTO-----EDFNKIMEYFEELKXANPPPGVKA----- 607

Qy 1041 DGQVLASRFWAYHKPLKNSQDYREALRAARELANITADLRKUPGTDPAFEVFPYTITNV 1100

Db 608 -----ELAGE-----SYLNVLVDNL 622

Qy 1101 FYEOLYTLPEGLPMLSCLLV-----PTPAVSCLLGLDLRSGLLNLSIVMLVDTVGF 1155

Db 623 VNEELKGISITGITVIIVLIPAIFRRTRTAVAM-----IMPFLGALWTIGY 669

Qy 1156 MALWDISYNVAVSLNLYS-AVGMSVEFVSHITSRPAISTKPTWLERAKE-----ATI 1206

Db 670 MGLAGIPFSQ-TLAGVVSMIVGLGDVGMHITHRF-----LEELKEGNKTPIITAME 720

Qy 1207 SMGSAPVAGVAMTWLPGILVIGLAKAQIOIPFRMLMLITLGLLHGLVPLPVILSYVG 1266

Db 721 SVGPGLIG-ALTTAGGFLAALTGLRTAIHDVG-----KVLAVG-----IFFSMLAAYL- 768

Qy 1267 PDVNPAL 1273

Db 769 --VTPAI 773

RESULT 37
T33156
hypothetical protein F56H1.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T33156
R;Graves, T.
submitted to the EMBL Data Library, May 1998
A;Description: The sequence of C. elegans cosmid F56H1.
A;Reference number: Z21293
A;Accession: T33156
A;Status: preliminary;
A;Molecule type: DNA
A;Residues: 1-969 <GRA>
A;Cross-references: UNIPROT:O76372; UNIPARC:UPI000017BA34; EMBL:AF067618; PIDN:AAC19198.
C;Experimental source: strain Bristol N2; clone F56H1
C;Genetics:

```
Db      880 AEHRIEATSGVGPVTLAATSTFLAGSM--LPALTHAFYQVGFLVWISITSWTSTP 937
      QY      1239 FFRNLNLT 1248
      Db      938 FYLPMLSLT 947

RESULT 38
A:75018
transport protein PAB1101 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: A75018
R:anonymous, Genoscope
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: A75018
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-746 <KAW>
A:Cross-references: UNIPROT:Q9U29; UNIPARC:UPI00003456A; GB:AJ248288; GB:AL096836; NID
A:Experimental source: strain Orsay
C:Genetics:
C:Superfamily: Archaeoglobus conserved hypothetical protein AF0459

Query Match      2.5%; Score 170; DB 2; Length 746;
Best Local Similarity 19.1%; Pred. No. 0.0091;
Matches 146; Conservative 120; Mismatches 290; Indels 210; Gaps 32;

QY      550 VPFPLAIGYKGDYSEAEALMTSLNYPAGDPLRLAQAALWEAEFLMEAFOR--- 606
Db      134 MLPPEARNLISGDYS---STIIIVNLN-----REKNQKALVRVYNDIEREIRTFP 183
QY      607 -----MAGMFQVTFPAERSLDEINRTTAEDLPFPATSYIVIFLYISALGYSYSSSRV 660
Db      184 QGVEVLTGDLGTYKILELQNDNRNTWA----ISGIIVVILLYF-----YKSPIRM 233
QY      661 MYDSKATLGGVAVVLGAVMAAMGPFYSYLGRSSILVILQVVPFLVLSVGAD-NIFIFVL 719
Db      234 LVPL-----VPLIFGVVW-TLGPMLGLPILD-IATTVGAMIIGIDYGVHVTR 283
QY      720 EYQLRRPCEPREVHIGRALGVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLA 779
Db      284 YTEE--RGKGKSLSEEAEEAIAETGKALLGNALTITAGFLALSISILPSLKLSSLVWVG 341
QY      780 VILDFELQMSAFVALLSLDSKRQEAASRLDVCCKVQPELPPQCGEGILLGPFQKAYAPF 839
Db      342 LGLAALNAVITPALAILSEEFREK-----VMKKEIIAIGG-----KSKISF 385
QY      840 LHLWITRGV-----VLLFLFALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLNRY 893
Db      386 IFSMLGKAIKRSPTALLIALIISGVSLYGASKITTEVRLEKMIPTDL----- 433
QY      894 FEVGAPVYFTTLGYNPSSBAGNAICSSAGCNPFSTQKI-QYATFEFQSYLAIPASS 952
Db      434 ----PEIQALSDIRSEFGQDEVTILIKADDVDRDPTLVLRDILRFEIREIKADSY----- 482
QY      953 WVDFFIDWLTSPSCCRLYISGPNKDFCPTSTVNSLNCNCSITMGSVRPSVEQPHKYL 1012
Db      483 -----INNV-----FEQSIADVVIOKYGI 503
QY      1013 PWFNLDRPNIKCPKGGIAAYSTSVNLT---SDGVLA-----SRFMAHYKPLKNSQDYT 1063
Db      504 P---QDKEKISEATEGSSLSVSDYSMTI IKLGNFMGVTOSEFNRIMEYFERAIQRADFP 560
QY      1064 EALRAABELAANITADLRKVPCTDPAFEVFPYITINVFQYLTILPEGLFMLSCLVPT 1123
Db      561 PGKRV--ELAGE-----SYLNVYLNLDVNEELGKISTIGTGLIVVWVFAL 603
QY      1124 FAVSCLLLGLDLSRGLNLLLSIVMLIVDTVGFMALMDISYNAV-----SLINLVS-AV 1175
```

```
Db      604 FR-----RPTVSIAMIMPFLG--ALMTIGYMGLAGIPFSQTLGAVVSMIV 647
      QY      1176 GMSVEFVSHITRSFAISTKPTWLERAKE-----ATTSMGSAVPAGVATNLPGLIVL 1227
      Db      648 GLGVDYGMHITHRF-----DEELNEGNTPIITALESVGPGIFVG-ALTAGGP--L 696
      QY      1228 GLAKAQLIQIFFRRLNLLITLLGLLHGLVFLPVLVSIVGPDVNPAL 1273
      Db      697 ALLSAQLTAIHDFGRVLAVG-----IFASMFAYL---VTPAI 731

RESULT 39
B81282
probable integral membrane protein Cj1373 [imported] - Campylobacter jejuni (strain NCTC
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: B81282
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chiller
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrel
Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hy
A:Reference number: B81250; MUID:20150912; PMID:10688204
A:Accession: B81282
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-823 <PAR>
A:Cross-references: UNIPROT:Q9PWS7; UNIPARC:UPI00000C1EE7; GB:AL1139078; GB:AL111168; NII
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: Cj1373

Query Match      2.2%; Score 154; DB 2; Length 823;
Best Local Similarity 18.3%; Pred. No. 0.014;
Matches 168; Conservative 153; Mismatches 351; Indels 244; Gaps 34;

QY      393 EKAFHDQHGFPPFTNQVILTAPNRSSRYRVDLSLLGPNFSGILDLDLL--LELLEQER 450
Db      76 EKPFSENELAKLKKLHELEKAP-----LVERVFS-IINAPLQSSQNTDLKEL 123
QY      451 LRHL-QVWSPEAQRNISLDICYAPLNPDTSLDYDCCINSLLQYFONNRTLLLTANQTL 509
Db      124 LKNIPNIESQDINRTKAQNEILNSPFYKNNIISKDKVTGLIILYKPTVY-----NNLI 178
QY      510 MGQTSQVDMKDHFLYCANAPLTFKDGITALCALSCMADYGA PVFPFLAIGGYKGYSEABA 569
Db      179 EKRDLATDEKE-----KQIRLAI-----KEHQDKQK 205
QY      570 LIMFTSLN-----NYPAGDPLRLAQAALWEAEFLMEARFQRMAGMFQVTFFTAERSLE 622
Db      206 VITKQSLDTIKSIVRNYEKNDTL-----YLGGVSMIA----- 238
QY      623 DEINRTTAEDLPFPATSYIVIFLYISLAL-GSYSSWSRVMSVDSKATLGLGGVAVVLGAVM 681
Db      239 DDMIAIYKSDLVLVGVS--LVFL-LGLALYFFRSWRVFLP-----LFICFISLS 286
QY      682 AAMGFFSYLGRSSILV---ILQVDFVLVSGADNIFIVLEYQRLPRRPGSPREHVIGR 738
Db      287 AASGVFALLNQITVISSNYVALVLIITLSVVVHLITHFIBSTQNHKPAKVE--RIVLET 344
QY      739 ALGRVAPMLLCSLSEACFFLGALTPM-PAVRTFALTS---GLAVILDPLLOMSAFVAL 794
Db      345 LLAKANPS-LVAIVTMTIGFFSLILSNIEPIIKLGIMMSIGLGLALISYFLASTILVLL 403
QY      795 LSLDSKRQEAASRLDVCCKVQPELPPQCGEGILLGFFQKAYAPFLHLWITRGVLLAFL 854
Db      404 KPNYHKKEKFNLLAFCAK-----TALDPKRRMIYISVLAIL 444
QY      855 ALFGVSLYSMSCHISVGLDQELALPKDSYLLDYFLFNRYFEVGA PVYFTTIGYNFSSA 914
Db      445 ALIGISKLRV-----ENSFV-----NYFKDGSBI-----KK 470
QY      915 GMAICSSAGCNPFSTQKIQYATFPP-----EQSYLAIPASSWVD-- 955
```


QY	665	KATLGAGVAVVLGVAAAMGFFSYLGRSSVLTVQVFFVLVSVG--ADNIFIFVLEYQ	722
Db	365	RATL-----IPVIAPIISIIIGTFAGMYVLGFSINLLTLFLGLVLAIGIWDDAIVLENVE	419
QY	723	R-LPRRPGEPREVHIGRALGV-----APSMILCSLSEAICFFLPGALTPMPAVRTEALTS	776
Db	420	RIMTTEKLSPRKAAT-KAMGEVTGVIAIVLVCAVFIPIVAFMGLGEM--YKQFAVTI	476
QY	777	GLAVILDFLLQMSAFVALLSLDSKRQEASRLDVCCVKPQBLPPPGQGEGLL-LGFFQKA	835
Db	477	AISVTLGSLVALTLPALCALILK-----PGHEPILPRIFNRA	516
QY	836	YAPFLLHWITRGV-----VLLFLALFGVSLYSMCHISVGL-----DQ-----	873
Db	517	PER-VTSGYTRGVRFELKRATIGLIIFAGLLGSTYLLFERVPGSLLPDEDQGLFGVAVL	575
QY	874	--ELALPKDSYLLDY-----FLELNRYFEVGCAPVYFV-----	903
Db	576	PPAASLERTTVVLQDVSENIKNRPADVNVFAVGFDLLSGGLKTSAGTMFIMLKDWKERT	635
QY	904	-----TTLGYNFSSEAGMNAICSSAGCNNSFT-----OKIOYA	937
Db	636	TPDADARNLPRTIMGWAGIKOGWLAFNPPFIMGSLTTGGPELVQDRTGGVESLTOA	695
QY	938	TEFPPOSYLAIPASSWDDDFIDWLTPSSCCRLYISGPNKDKCPSTVNSLCLNCKMSIT	997
Db	696	TKLITEAANKPELOQVTRTFDPYPOYDIQL-----DREKAKMGV-PINSVFTAMQAT	749
QY	998	MGSVRPSVEQFHKY-----LPWFLNDRPNIK-----CPKGG--LAAYSTSVNLT	1039
Db	750	FGSL--YVNDFTLYGRNYQVNIQSEAEFRDGDGLKHVPVRADSGSMPLDALVTVKRIV	807
QY	1040	SDGQVLASRFMAYH--KPLKN--SQDYT--EALRAARELAANTADLRKVPGTDPAFE--	1091
Db	808	GPDL--ERFNAPNAKVTGNPAGYTSGDAIKAMQEVAAQVLPQGYQIAWTGSAYQEV	865
QY	1092	-----VFPYTTNVFEQYLTILPEGLFMLSCLVPTFAVNSCLLGLDLRSG	1138
Db	866	TSGTGSQAMIFGLIMVFLILAAQYERW--SLP-----LAVITAVPFAIPGALLATDLR-G	917
QY	1139	LNLLSIVMLVDTVGFMAWLDISYNAVSLINLVSAGVMSVFEVSHITRSPAISTKPTWL	1198
Db	918	LTV-----DYVFO-----IGLVTLIGLAVKNAIILV-EFAV-----L	948
QY	1199	ER-----AKEATISMGSAVFAGVAMTNLPGLIL-VLGLA	1230
Db	949	ERESGKSAIEAASAAARLPRPIVMTSLATPILGWPLA	986
RESULT 42			
H84301			
hypothetical protein Vng1476c [imported] - Halobacterium sp. NRC-1			
C;Species: Halobacterium sp. NRC-1			
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004			
C;Accession: H84301			
R;NG, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jable			
Jung, K.H.; Alam, M.; Freitas, T.			
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000			
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li			
A;Title: Genome sequence of Halobacterium species NRC-1.			
A;Reference number: A84160; MUID:20504483; PMID:11016950			
A;Accession: H84301			
A;Status: preliminary			
A;Molecule type: DNA			
A;Residues: 1-746 <STO>			
A;Cross-references: UNIPROT:Q9HPT7; UNIPARC:UP100000638D0; GB:AE004437; NID:g10580974; H			
C;Genetics:			
A;Gene: VNG1476C			
Query Match			
Best Local Similarity 2.1%; Score 148.5; DB 2; Length 746;			
Best Local Similarity 21.6%; Pred. No. 0.029;			

Matches	175;	Conservative	99;	Mismatches	288;	Indels	247;	Gaps	41;
Qy	586	LAQAKLWEAFLEEMRAFORRMAGMFQVFTFAERSLEDRNRTABDLP-----IF	636						
Db	71	LIRGDIPTESAIBADRYDRRMSQTODV--YAVRSPADQV-RAAYGRIPDSKIESVIG	127						
Qy	637	ATSYIVIFLYISIALG-----SYSSWSRVNVDKAT-----	667						
Db	128	SPDYTTIQATDLDLQQOORPIYTEAVSTSDWARFPASVTAIVTGDAAPSAQLSTVIOQ	187						
Qy	668	-----LGLGG-----VAVVLGAVMAAMGPFYSYLGRISLVLQV	701						
Db	188	STQKLLGLAVGLMVALFFLFRGVRLRLLPVAVFVG-VIYTFGAIGYGGIPNSTMTSAV	246						
Qy	702	VPELVLSVGADNIFIFVLEY-QRLPRRPGEPREVHIGRALGRVAPSMILCSLSEAICPFL	760						
Db	247	FPILI-GLGIDYSVQFHRYEESELETHP--PREA-LPNALCGVGPVPVLIANLAAALGF--	300						
Qy	761	GAL---TPMPAVRTFALTSLAVILDPLQMSAFVALLSLDSKR-----	801						
Db	301	GATVSTSSPAIIWFAQTSILGVLLTYAAGILVLLAGLTIIYIRRWGDDSPAADGGDPEA	360						
Qy	802	-----QASRLDVCCVKPOELPPPGQGEGLLGFQKAYAPFLHWHITRGVLLFLAL	856						
Db	361	AHAEVDEDSRIGLV-----GSTLGRTSRIILAS--HPMT--VLAIALVL	399						
Qy	857	FGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTILGYNFSBAGM	916						
Db	400	SGVGFQVASSLGLTADTEEFIPQD-----LPAVDL-----QQFRSQTG-	438						
Qy	917	NAICSSAGCNNSFTQKIQYATFEFPQSYLAIPAS-SWDDDFIDWL-----TPSS	965						
Db	439	-----GGTSTSVDLV-----TGTHLKHPETLRWMDGLNDVATQAPLVEGVDTPAG	484						
Qy	966	CCRLIYISGPNKDKCPSTVNSLCLNCKMSITMGSVRPSVEQFHKLPWFNLDRPNIKCP	1025						
Db	485	AVKQYNGE-----IPRTSGVERVLAQM-----PTEQORY--YNDGRAHI----	524						
Qy	1026	KGGLAAYSTVNLTSDQVLASRFMAYHKPLKNSQ---DYTEALRAARELAANITADLRK	1082						
Db	525	-----TVASPDQMTT-QGVLSF-----LENSQESLDYSNP-----PPGVTAIE---	560						
Qy	1083	VPGTDPAPFEPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVNSCLLGL---DLRSLG	1139						
Db	561	VTGTSSI--AVPSIIDQINSRNTTT---GLGILFV-----FA---LLGLYYRDLVKAV	605						
Qy	1140	LNLLSIVMLVDTVGFMAWLDISYNAVSLINLVSAGVMSVFEVSHITRSPAISTKPTW--	1197						
Db	606	APLVPVMFEVIGWQNIYMLALDIPVSPIGASLGAMTVGIGAEYTIIVMERYEERQHNQY	665						
Qy	1198	-LERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQIPFFRLNLLITLLGLLHGLV	1256						
Db	666	RLDVAEIAASRVGKAITVS-GMTTVFGFSALISPPFVLSFGF-----LTVGVI	714						
Qy	1257	FLPVTLXVGGDVNPDALAEQKRAEAVA	1285						
Db	715	FLTLIASLA--TLPTPLVLDGLAERVA	741						
RESULT 43									
AC0423									
multidrug efflux protein [imported] - Yersinia pestis (strain CO92)									
C;Species: Yersinia pestis									
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004									
C;Accession: AC0423									
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;									
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;									
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.;									
Nature 413, 523-527, 2001									
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.									
A;Reference number: AB0001; MUID:21470413; PMID:11586360									
A;Accession: AC0423									
A;Status: preliminary									

A:Molecule type: DNA
A:Residues: 1-1041 <KUR>
A:Cross-references: UNIPROT:Q8ZBD4; UNIPARC:UPI00000DC645; GB:AL590842; PIDN:CAC92711.1;
C:Genetics:
A:Gene: YPO3482
C:Superfamily: acriflavin resistance protein

Query Match 2.1%; Score 146; DB 2; Length 1041;
Best Local Similarity 20.2%; Pred. No. 0.068;
Matches 168; Conservative 120; Mismatches 305; Indels 246; Gaps 41;

QY 562 KDYSAEALIMTFSLNNYPAGDRL-----AQAQWEEAFLEEMRAFORRMAGMFOVTT 616
DB 274 EYIS-----VTRLNHGAAGIAVLMAGALATAEARKAKAEFEULPDGKIAPP 327
QY 617 AE-----RSLEDEINRTAEDLPFATYSIVIFLYISLALGSYSWSRVMVDKATLIG 671
DB 328 KDSTDPIKYSVEEVVKTLIE-----ALLVIVIMYIFLQ-----NIRATL-IP 369
QY 672 GVAVVLGAVNAAMGPFYSYIGIRSSVLQVPLVLSVG--ADNIPFVLEYORLPRRG 729
DB 370 AIAY-----PVLLGTGVLAIFGYSINTLTLCMVLSIGLLVDDAIVVVENVERVMREDN 425
QY 730 EPREVHIGRALGVAPSMCLLSSEAICE-----FLGALTPMPAVRTFALTSGLAIVILDF 785
DB 426 LPPEATEKSMSEIASALIGIALVLSAVFLPMAFFGATGV-IYRQFSITIVSAMALSVL 484
QY 786 LQMSAFVALSLDSKRQEAERLDVCCVKPQBLPPPGQEGLLGF-----FOKAYAPP 839
DB 485 VALTLTPALCA-----TFLKPNKHP--SEHGFGFGRYDRMQRYESL 528
QY 840 LLHWITRGV-VLLFLALFGVSLYNCHISVGLDQELALPKD-----SYLLDYFLFLNR 892
DB 529 VGRVHRSRLYLLIYAVLIGV-----MCVLFIRLPTGFLPTEQDQGVDMVQYTLPAGATSK 584
QY 893 YEEVGAPV--YFV-----ITLGNFYS---SEAGMNAIC-----SSAGCNFPST 931
DB 585 TWEVSKAVENYFMTQEKONTKAVFTISGFGSGSQONAGMFIALKHWRDRPSEN---- 640
QY 932 QKIQVATEPEQSYLAIAPASSWVDVDFIDMLTPSSCCRL-----LGIRSSLVI----- 698
DB 641 ----TATAIADRAMKAL--SSIRDAQIPMTTPAVDGLQSGNGFTPELQATGDTSEQLL 694
QY 970 -----YISGNKDKFCPST-VNSLNLKN-----CMSITWGSVRPSVEQPHKYL 1012
DB 695 TLRDQLISKANKDPIILASVRANTLQMPOLQVDINDKVAALGLSISDVNATLS-----A 749
QY 1013 PW---FLN---DRPNIK--CPKGLAAYS-----TSVNLTSDCQVLSRPMAY 1052
DB 750 AMGGTYINDFIDGRVKRYMQGDVDTRSKPEDLNQWFMVRSDDAMTSFSAFATTWYIG 809
QY 1053 HKPLK--NSQDYTE-----ALRAARELAANITADLRKVPGTDPAPFVFPYTI 1097
DB 810 PTLRSRYNGQTSYEYEQQAASGSSSGSTAMDQMEKLA-----ELPGTSYAWSGLSVQ- 861
QY 1098 TNVFEQVLTILPEGLFMLSCL-----YPTFVAVSCLLGL-----DLRSLGLNL 1142
DB 862 ERLASQALSLVAISLVVFLCLAALYBSWSVPFSVMVVIPLIGIICAVAAATLR-GLEND 920
QY 1143 LSVIMLVDTVGFMAWDISYNAVSLINLVSAVGM-----SV 1179
DB 921 IYFQVALLTTLGL-----ASKNAILVPEAAAYLRGEPLVVAALQGAATRLRPILMTSL 975
QY 1180 EPVSHITRSFATSKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLAKAQLIQIF 1238
DB 976 AFIAVG-MPLAMST-----GAGANKRISIGSIIGTTLTATVLAFFVPLFPVLRVP 1028

RESULT 44

H69780

antibiotic transport-associated protein homolog ydfJ - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C:Accession: H69780
R:Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte-
C:Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch-
A: Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.;
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Frit, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Galle-
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maugda, S.; Maues
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli-
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Sero-
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, F.; Tognoni, K.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, F.
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: H69780
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-724 <KUN>
A:Cross-references: UNIPROT:P96687; UNIPARC:UPI000005FFBF; GB:Z99106; GB:AL009126; NID-
A:Experimental source: strain 168
C:Genetics:
A:Gene: ydfJ

Query Match 2.1%; Score 143; DB 2; Length 724;
Best Local Similarity 20.2%; Pred. No. 0.068;
Matches 137; Conservative 101; Mismatches 249; Indels 192; Gaps 30;

QY 666 ATLGIGGVAVVLGAVNA-----AMGPFYSY-----LGIRSSLVI----- 698
DB 169 ARMEIGGVSEIVGIILAFVWLAITFGSLIAGLPIALTALIGLVSIGLVLIGTQVFDIAS 228
QY 699 --LQVVPVLVLSVGAD-NIFPVLSEYORLPRPGPREVHIGRALGVAPSMCLLSLSA 755
DB 229 VLSLAGMIGLAVGIDYALFIPTKTRQFLGE--GTQKNESIAARAVGTAGSAVFAGLTVI 286
QY 756 ICFELGALTMPAVRTFALTSGLAIVLDFLQMSAFVALLSDSKRQEAERLDVCCVKP 815
DB 287 VALCGLTVNIPFMSAMGLTAGSLVMAVLASITLVPVLSIAGKRM-----IPKS 337
QY 816 QELPPPGQEGLLGFPQKAYAPFLHWTIRGVLLLLFLALFGVSLYNCHISVGLDQEL 875
DB 338 NKKIEKQSTETNVMGRF-----VTKNPIML-----SVCSILLILI--VI 373
QY 876 ALPKDSYLLDPLFLNRYFEVGPVYFTLTYGNFSSBAGMNAICSSAGCNFSTQKIQ 935
DB 374 SIP-----SMHLGLP-----DAGMK----- 390
QY 936 YATEPEQSYLAIAPASSWVDVDFIDMLTPSSCCRLYISGNKDKFCPSTVNSLNLKNCMS 995
DB 391 -AKONPDRAYDLLAEGFEGENGQLT--IVADATNATENKAAADAVKEIKGLDHVAS 447
QY 996 ITMGSVRPSVEQPHKYLFWLNDPRNIKCPKGLAAYSTSVNLTSDCQVLSRPMAYHKP 1055
DB 448 VT-----PAM-----PN-----KGNFALITVVPETGPNVDTKDLVDHVS 484
QY 1056 L--KNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYTIINVFQVLTILPEGL 1113
DB 485 LSDKNGVD-----LLVTGSTAVNIDISDR-----LNDALPVPFVLIIVG-FAFVLLTIVPFRSL 535
QY 1114 FMLSCLVPTFAVSCLLGLDLRSLGLNLLSIVMLVD--TVGFMAWD-----ISYNAVS 1167
DB 536 -----LVPLVAVAGFMULTMTATIGI-----CVFVLQGNLIDFPKIKPEKGPILAPLPL 584
QY 1168 LINLVSAGVM--SVFVSHITRSFAISTKP-----TWLERAKEATISMG-----SAV 1212
DB 585 SIGILFGLAMDYQVFLVSRMRBEYVKTNNPVQAIQAGLKHSGVPVTAAGLIMI FVPAGFI 644
QY 1213 FAGVAMTNLPGLIVLGLAKAQLIQIFPFRMLMLITLIGLL-HGLVFLPVLVSIVGPDVN- 1270
DB 645 PAGEASIKANG---LALSFGVLDFDAFIVRMTLIPSMKLMGNAAWYLPKWLDKIIPNVDI 701

Qy 1271 -----PALALEQKR 1279
Db 702 EGHQLTKEIQEIDHEQKK 720

RESULT 45
AF0561
A:Title: acriflavin resistance protein B [imported] - Salmonella enterica subsp. enterica serovar
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: This species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AF0561
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AF0561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <PAR>
A:Cross-references: UNIPARC:UPI000005A1FB; GB:ALU513382; PIDN:CAD04960.1; PID:g16501747;
C:Genetics:
A:Gene: STY0519
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 141.5; DB 2; Length 1049;
Best Local Similarity 18.1%; Pred. No. 0.14;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;

Qy 624 EINRTAEDLPFATSYIVIFLYISALGSSWSRVVDSKATL--GLGGVAVVLG--A 679
Db 339 EVKTLVE-----AILVFLVMYLFQNF-----RATLPIATVPVVLGTFA 381

Qy 680 VMAAMGFFSYLGIRSLVILQVVPFLVSVG--ADNIFIVLEYQRLPRRPGEPREVIHG 737
Db 382 VLAAGF-----SINTLTMF--MVLAIGLLVDDAIIVVENVRVMTTEGLPPKEATR 432

Qy 738 RALGR-----VAPSMILCSLSEAIQFELGALTTPMAVTRFALTSLGLAVILDFLLQMSAFV 792
Db 433 KSGQIQGALGVAMVLSAVFTPMAFGGSTGAI--YRQFSITIVSMALSVLVALITLP 490

Qy 793 ALLSLDSKRQEARSLDVCCCKVPQELPPPGQEGLLGFFQKAYAPFLLHWI----- 844
Db 491 ALCA-----TWLKFVAKGDHGGKGGFFGNFLFDKSTHHYTDVSGNILR 536

Qy 845 TRGVLLLLFLALP-----GVSLYSNMCHISVGLDQELALPKDSYLL 884
Db 537 STGRYLLLLIIVVGMAVLFVRLPSSFLPDEQGVFL--TMVQLPAGATQERTQKVLDEVT 595

Qy 885 DYEL-----FLNR----- 892
Db 596 DYLYNKEKANVESVFAVNGFGFAGRGQNTGIAFVSLKOWADRPGEKVKYALTQRTAF 655

Qy 893 -----YFEVGAPVYFVTTLLGYNFSSEAGKNAICSSAGCNFSFTQ----- 932
Db 656 SQIKDAMVFAFNLPAIVELGTATGDFE-----LIDQAGLGHKLTQARNQLGEVAK 708

Qy 933 -----KIQYATEPEQSYLAIPAS-----SWVDFTDWLTP 963
Db 709 YPDLLGVVRPNLEDTQPKIDIDQEKAAQALGVSISDINTTLGAANGSGSYVNDIFD--R 765

Qy 964 SSCRLYISGNPKDKFCPSVNSLNCINCKMSITMGSVRPSVQPHKLPFWLNDRNPIK 1023
Db 766 GRVKVYVMEAKYRMLPDDIND-----WYVRGSDGQM 798

Qy 1024 CPKGGAAYST-----SVNLTSDGQVLAIRFMAHYHKPLKNSQDYTEALRAARELANIT 1077
Db 799 VP---FSAFSSRWYEGSPRLRYNGLPSMELLQQAAPGKST---GEAMAMBEELAKLP 852

Qy 1078 ADLRKVPGTDPAPEVPPYTTITNVFBOYLT-----ILPEGLFMLSCL----- 1120
Db 853 SGI-----GYDWTGMSYQERLSGNQAPALYALISLIVVFLCLAALVESMSIPP 899

Qy 1121 -----VPTFAVSCLLGLDLRSGLLMLLSIVMLVDTVTFMAMLDISYNVSLIN----- 1170
Db 900 SVMLVVPLGVIGALLAA--TFRGLTNDVYFQVGLLTITGLSA-----KNAILIVEFAKOL 952

Qy 1171 -----LVSAYGMSVEFFVSHITRSPALSTKPTWLERAKEATISMGSAVPAGVAMT 1219
Db 953 MDKEGKGLVEAMLEAVRMRLRPILMTSLAPMLGVP-----LVISSG-----AGSGAQ 1000

Qy 1220 NLPQILVL--GLAKAQLIQIPFF 1239
Db 1001 NAVGTGVLGGWVTATVLAIFF 1021

RESULT 46
D85757
A:Title: efflux pump Z2508 [imported] - Escherichia coli (strain O157:H7, substrain BDJ9
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: D85757
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1027 <STO>
A:Cross-references: UNIPROT:Q8X7E2; UNIPARC:UPI000000D0A77; GB:AE005174; NID:g12515493;
A:Experimental source: strain O157:H7, substrain BDJ933
C:Genetics:
A:Gene: Z2508
C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 140.5; DB 2; Length 1027;
Best Local Similarity 18.3%; Pred. No. 0.16;
Matches 149; Conservative 125; Mismatches 261; Indels 281; Gaps 38;

Qy 576 LNNTYA-----GDPRLAQAKLWBEAFLEENRAFQRMAGMFQVTFTHAERSLEDBIN 626
Db 280 LNHPSAGVAVMLSPGANALNTATLVKDIAE----FQNNPQGYDIAYPKDST---BFI 332

Qy 627 RTTAEADLPFATSYIVIFLYISALGSSWSRVVDSKATLGLGGVAVVLGAVMAANGF 686
Db 333 KISVED--VIQTLFEAIVLVVCMYLFQNLRLATLIPALA-----VPVLLGTGFGVJAL 384

Qy 687 PSYLGIRSLVILQVVPFLVSVG--ADNIFIVLEYQRLPRRPGEPREVIHGRALGRVA 744
Db 385 FGY-----SINTLTFA--MVLAIGLLVDDAIIVVENVERIMRDKGLPAREATEKSMGEIS 438

Qy 745 PSMLICSLSEAIQF-----FLGALTTPMPAVTRFALTSLGLAVILDFLLQMSAFVLLSLSOK 800
Db 439 GALVAIALVLSAVFLPMAFFGSGTGV--YRQFSITIIISAMLSVVVALTLTPAL----- 491

Qy 801 RQEARLDVCCCKVPQELPPPGQEGLLGFFQKAYAPFLLHWITRGVLLFLALFGVS 860
Db 492 -----CGSVLQHVPPHKG---FFGAPDFRFRRTEDKY--QRGVIYVLRRAARTMG 537

Qy 861 LYSNMCHISVGL-----DQ-----ELALPK-----DSYLLDYFLFLNR 892
Db 538 LYLVLGGGHALMMWKLPGSFLPTEQGEIMVQYTLTPAGATAARTAEVNRQIVDWF--INE 596

Qy 893 YFEVGAPVYFVTTLLGYNFSSEAGMN----- 917
Db 597 --KANTDVIYF--IVDGFSGSG--SQNTGMAFVSLKNWSQRKGAENTAQAIALRATKELGTI 652

Qy 918 -----AICSSAGCNFSF-----TOKIOYATEPFE----- 942
Db 653 RDATVFAMTPPAVDGLGQSGNGFTFELLANGGTDRETLQMRNQLTEKANQSPELHSVRAN 712

QY 943 -----QSYLAIP-----ASSW-----VDDFIDWLTSPSSCCRLYISGP 974
Db 713 DLPMPQLQVDIDSNKAVSLGSLNDVDTLSSAWGGTYVDFID---RGRVKKVYIQGD 769
QY 975 NKDXFCSTVNSLNLCKNCSITGWSVRPSVEQPHKYPWFLNDRPNIKCPKGLAAAYST 1034
Db 770 SEFSAFSDLGK-----WVVRGSDNAMTP-----FSAFAT 799
QY 1035 SVNLTSDGOVLA-SRFMAYHKPLKNSQDYT---EALRAARELAANITA-----DLRK 1082
Db 800 TRMLYGPRLVRYNGSAAYEIQENAGTGFSSGDAMTKWEELANSLPAGTTWAMSGLSLQ 859
QY 1083 VPGTDPAFEPVPIITNVF-----YEQY-----LTILPEGLFMLSCLVPTFAVSCLL 1130
Db 860 KLASGQALSLYANSILVVFVCLAAALYESWSVPFVSVILVPLG-----L 902
QY 1131 LGLDLRGLNLNLISIV---MLVDTVGFMAWDISYNVSLINLVNSAVGMSVVFVSHITR 1187
Db 903 LGAALAAWRDLNDVVFQVALLTTIGLS-----SKNALIVFEAEAA-----VAE 948
QY 1188 SFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPG 1223
Db 949 GYSLSRAALRAAQTRLRPIIMTSLAF--IAGGNAAG 982

RESULT 47
H90861
probable multidrug-efflux transport protein ECs1864 [imported] - Escherichia coli (strain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90861
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaka, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90861
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1027 <HAY>
A;Cross-references: UNIPROT:Q8X7E2, UNIPARC:UPI00000DOA77; GB:BA000007; PIDN:BA035287.1;
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs1864
C;Superfamily: acriflavin resistance protein

Query Match 2.08; Score 140.5; DB 2; Length 1027;
Best Local Similarity 18.3%; Pred. No. 0.16;
Matches 149; Conservative 125; Mismatches 261; Indels 281; Gaps 38;
QY 576 LNNYPA-----GDPLRLAQAKLWEAFLEENRAFQRRMAGMPOVTTFAERSLEDEIN 626
Db 280 LNHGPSAGVAVMLSPGANALNTATLVKDKTAE-----FORMNPGYDIAPKOST---EFI 332
QY 627 RTTAEDLPFIATSYIVIFLYISIALGSSYSSWRVWDSKATLGLGGVAVVLGAVMAAMGF 686
Db 333 KISVED--VIQTLFEAILVVCVMYFLQNLRLATLIPALA-----VPPVLLGTGFGVAL 384
QY 687 FSYLGIRSSLVILQVPPFLVLSVG--ADNIFIFVLEYQRLPRPQGPPEVHIGRALGRVA 744
Db 385 FGY-----SINTLTFLA-MVLAIGLAVDDDAIVVVENVERIMRDKGLPAREATEKSNGEIS 438
QY 745 PSMLLCSLSRAICF---FLGALTTPPAVTEPALTSGSLAVILDPLQMSAFVALSLDSK 800
Db 439 GALVAIALVLSAVFLPMAFPGGSGTV--IYQFQSIITISAMLLSVVALLTUTPAL-----491
QY 801 QEASRLDVCCKVPQBLPPPGQEGILLGFFOKAYAPFLHWTIRGVWLLIFLALFGVS 860
Db 492 -----CGSLQHVHPHKKG---FFCAFDPRFYRTEDKY-QRGVIVVLERAAATWG 537
QY 861 LYSWCHISVGL-----DQ-----ELALPK-----DSVLLDYFLFLNR 892

Db 538 LYLVLGGGHALMMWKLPGSFLLPTDQGGIMVQYITLPAGATAARTAEVNRQIVDWFL--INS 596
 QY 893 YFEVGAPYFYVTTLTGYNFSSBAGMN----- 917
 Db 597 --KANTDIVP--TVDGFSFSG--SGQGTGMAFVSLKNWSORKGAENTAAQALRALRATKELGTI 652
 QY 918 -----AICSSACGNNFSP-----TKQIYATEPPE----- 942
 Db 653 RDATVFAMTPPAVDGLGQSGNGFTFELLANGGTDRETLQMRNQLIEKANQSPELHSHVRAN 712
 QY 943 -----QSYLAIP-----ASSW-----VDDFDLWLPSSCCRLYISGP 974
 Db 713 DLQPMPQLQVDIDSKNAVSLGSLNDVDTLTSAGGTYGVNDIFD---RGRVKKVIYIQGD 769
 QY 975 NKDKFCPPSTVNSLNLCKNCSITMGSVRPSVEQPHKLPFLNDRPNIKCKPCKGGLAAYST 1034
 Db 770 SEFRSPASDLCK-----YEOY-----LTLPSGLFMLSCLVPTPAVSCLL 1130
 QY 1035 SVNLTSDQGVLA--SRFMAYHKPLKNSQDYT--EALRAARELAANITA-----DIRK 1082
 Db 800 TRWLYGPERLVRVYNGSAAYEQENATGFSFGDAMTKMEELANSLPAGCTTMAWSGLSLQE 859
 QY 1083 VPGTDPAPEVPYITTVF-----YEOY-----LTLPSGLFMLSCLVPTPAVSCLL 1130
 Db 860 KLASQALSLYAVSLVLFVCLALYESWSPFVSILVPIGL-----L 902
 QY 1131 LGLDLRSGLLNLSTV---MILVDTVGFPMALWDISYNAVSLINLVSAGMSVEFVSHITR 1187
 Db 903 LGAALAAWRDLNDVYFQVALLTIGLS-----SKNAILIVEFAEAA-----VAE 948
 QY 1188 SFALSTKPTWLERAKETATISMGSAVPAGVAMTNLPG 1223
 Db 949 GYLSLRALRAAQTRLRPIIMTSLAF--IAGGNAAG 982

RESULT 48
 S39630
 Multidrug-efflux transport protein B - *Pseudomonas aeruginosa*
 N:Alternate names: multidrug resistance protein B
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S39630
 R:Poole, K.; Heinrichs, D.E.; Neshat, S.
 Mol. Microbiol. 10, 529-544, 1993
 A:Title: Cloning and sequence analysis of an EnvCD homologue in *Pseudomonas aeruginosa*:
 A:Reference number: S39629; MUID:95058196; PMID:7968531
 A:Accession: S39630
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1046 <POO>
 A:Cross-references: UNIPROT:P52002; UNIPARC:UPI000016PC9A; GB:L11616; NID:9438852; PIDN:
 A:Experimental source: strain CD10
 C:Genetics:
 A:Gene: mexB
 C:Function:
 A:Description: probably involved in secretion of the siderophore pyoverdine
 A:Note: inducible under conditions of iron limitation
 C:Superfamily: acriflavin resistance protein
 C:Keywords: transmembrane protein; transport protein
 F:14-30/Domain: transmembrane #status predicted <TM1>
 F:343-359/Domain: transmembrane #status predicted <TM2>
 F:370-386/Domain: transmembrane #status predicted <TM3>
 F:397-413/Domain: transmembrane #status predicted <TM4>
 F:442-458/Domain: transmembrane #status predicted <TM5>
 F:472-488/Domain: transmembrane #status predicted <TM6>
 F:541-557/Domain: transmembrane #status predicted <TM7>
 F:874-890/Domain: transmembrane #status predicted <TM8>
 F:898-914/Domain: transmembrane #status predicted <TM9>
 F:973-989/Domain: transmembrane #status predicted <TM10>
 F:1014-1030/Domain: transmembrane #status predicted <TM11>

Query Local 2.0%; Score 140.5; DB 2; Length 1046;
 Best Match Similarity 20.5%; Pred. No. 0.17;

Matches	187;	Conservative	126;	Mismatch	305;	Indels	293;	Gaps	49;
Qy	506	NOTLMGOT--SQVDWKDHFYICANAPLTFKDGDTALALSCWADYGAPVFPFLAIGGYKGD	563						
Dd	231	NATIIKTRLOTAEQENILKVN-----PDGSOVELKDQADV-----LG---GQD	274						
Qy	564	YSBAEALIMTFSLNNVPA-----GDPRLAOKLWEE--AFLEEMRAFORMMAGHMQ	612						
Dd	275	YS-----INAFNGSPASGIAIKLATGANALDTAKAIRQTYTANLEPFPMQCKKVVYPD	328						
Qy	613	VTFATERSLEDEINRTAEDLPFATSYIVIFYISLALGSSYSSWSRVMVDGSKATL--GL	670						
Dd	329	TTPVASAIH-EVKTLGE-----AILLVFLVMYLFLONP-----RATLIPPI	370						
Qy	671	GGVAVVLG--AVMAAMGFFSYLGIRSLVILQVVPFLVLSVG--ADNIFIVFLEVORLPR	726						
Dd	371	AVPVLLGTGVLAAAGF-----SINTLTWFG-WVLAIGLLVDDAIWVENVERVMA	421						
Qy	727	RPG-EPREV--HIGRALGR-VAPSMLLCSLSAICFFLIGALTTPMAVTRFALTSLAVI	781						
Dd	422	EEGLSPREAAARKMGQTOGALVGIAVMLSAVFLPMAFFGGSTGVI--YRQFSITIVSAMA	479						
Qy	782	LDFFLOWSAFVALLSLDSKRQASRLDVCCL--VKPOELPPPGQCEGLGLLFFQKAYAP	838						
Dd	480	L-----SVIVALI-----LTPALCATMLKPIEKDGEHKGKGGFFGWFNMFLS	522						
Qy	839	FLHWTIRGVV-----LLAFALFGVSLYSMCHISVGL-----DQ-----ELALPKD	880						
Dd	523	-TTHGYERGVASILKHRAPVLLIIVVIVAGMIWMFTRIPTAFLPDEDOGVLFQAQVOTPPG	581						
Qy	881	S-----YLLDFELFLNRYFEVGAPVYFVTILGYNPSSE-----913							
Dd	582	SSAERTQVVDMSREYLLEK-----ESSVSSSVFTVTGFPAGRGQSGGMAFIMLKP	633						
Qy	914	-----AGNAI--CSSAGCNMFSTQKIQVATYEFPEQSYLAIPASSWDDFID-----959							
Dd	634	WEERPGENSVFELAKRAQMHFFSFKDAWVFA--FAPPSVLELGNATGDFLFDQDQAGVG	691						
Qy	960	-----WLTPSSCCRYIYISGPNKDKFCPTVN-----SINCLKNKMSI	996						
Dd	692	HEVLLQARNKEFLMLAAQNPALQVRPNMGMSDEPOYKLEIDDEKASALGVSLADINSTVSI	751						
Qy	997	TWGSV-----RPSVEQFHXYLPWFLENDRN-----IKCPKGGLA-----AYST	1034						
Dd	752	AWGSSYNDFIDRGVRKV-----YLQGRDARNPDDLKSKTVRNDKMGKVPFNAPAT	805						
Qy	1035	-----SVNLTSDGOVLASRFMAHYHKLNSQDYTEALRAARELAANITADLRKVPCTDP	1088						
Dd	806	GKWEVGSFKLERYNCVPAMEILGEPACPLSSGD--AMAAVEE---IVKQLPKGVG---	855						
Qy	1089	AFEVFPYITTVFVEQYIT-----ILPEGLFMLSCL-----LYPTFAVS	1127						
Dd	856	-----YSWTGLSYEERLSGQAPALYALSLLVFLCLAALYESWSIPFSVMLVPLGVI	909						
Qy	1128	CLLLGLDRSLGLNLLSIVMLILVDTPGFMAWDISYNVSLNL-----V	1172						
Dd	910	GALLATSMR-GLSNDVFFQVGLLTTIGUSA-----KNALLIVFAKELHEQCKGIVEAAI	963						
Qy	1173	SAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIYL-GLAK	1231						
Dd	964	EACRMRLAPIVMTSLAFILGVVP-----LAISTG-----AGSGSQAIGTGVIGGMVT	1011						
Qy	1232	AQLIQIFPFRRL	1242						
Dd	1012	ATVLAIFWVPL	1022						

RESULT 49
E83593
RND multidrug efflux transporter MexB PA0426 [imported] - *Pseudomonas aeruginosa* (strain
C/Species: *Pseudomonas aeruginosa*
C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C/Accession: E83593

R; Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, L.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A; Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen

A; Reference number: AB2950; MUID:20437337; PMID:10984043

A; Accession: E83593

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-1046 <STO>

A; Cross-references: UNIPROT:P52002; UNIPARC:UPI000012F022; GB:AE004479; GB:AE004091; NII

A; Experimental source: strain PA01

C; Genetics:

A; Gene: mexB; PA0426

C; Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 140.5; DB 2; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.17;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

C:\date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2000
C:\Accession: E83593


```
Db 133 QSNWILLANGLYSQQEYNTVFLSNYADQYIVDALKRLDGVGEARIIGERRYANRLMDP 192
Qy 460 E--AQRNISLQDICVAPLNPDN-----TSLYDCCTNSLLQYFQNNRTLLLTAN 506
Db 193 NRLASRLTAQVIDA-INEQNIQGVGIGQPSLPD-----QWYQ-----IDLQAR 239
Qy 507 QTLMGQTSQVDWKHFLYCANAPLTFKDTALALSCM--ADYGAPVF-PFLAIGYKGD 563
Db 240 GLRSEATEFADM-----VITGSDGTILKLDVGRABELGAENYSFLRFRGNEG-- 288
Qy 564 YSEAEALITSLNNYPAGDPLRAQAKLWEEAFLEEMRAFORRMAGM-----FQVT 614
Db 289 -----VGIGFPTPGSNA-----LEVSXAVTEMARLAQSPFGKQYQA 328
Qy 615 FTAERSLED---EINRTTAEDLPIPATSYIVIFLYISALGSYSWSRWVDSKATLGLG 671
Db 329 FDTTLFVBESLAELVVKTLIEALILV--LVIFEL-----QDMRTTLIP----- 370
Qy 672 GVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVG--ADNIFIFVLEYQRLPRPG 729
Db 371 -----VITTLTGITGFAIKVFGFSINTLTFLGLTLATGLVDDAILIIVENISRLIEQG 426
Qy 730 EPREV---HIGRALGRV-APSMILCSLSEACFFLGAITPMPAVRTPALTGLAVILDF 784
Db 427 NSPROASESMRELFGAVIATSLVMVFPVPAFFPG--TTGQIYKQFALTIAFSWVIST 484
Qy 785 LQMSAFVALLSLDSKROBASRLVCCCVKQPELPPQGGELLLGFFQKAYAPP--LLH 842
Db 485 FLALTLPSSLALLRRGQRP-----VLLLELALF-----GWLKGFHQVNRFLD 521
Qy 843 WITRG-----VLLLELALF-----G 858
Db 522 WRRRGYERSLNFKVRAIVLLEFLASIALTGWVLSVPTAFIPEDQGYFITIQGPEG 581
Qy 859 VSLYSMCHISVGLDOB-LALPK--DSYLDYELFLNRYFEVGAPYFVTTLGYNFSBAG 915
Db 582 VSLNYSNVMAQVEIKLPEVTGTFAGIGFSFGNSANSNG--VIFTLQWDERQPG 639
Qy 916 MNA---ICSSAGC-----NNPSF----- 930
Db 640 QSAPEIIGLAGVFSNITEARIFVSPPPIDGLSGFGQFELQDRAGNTGLNLLQVMG 699
Qy 931 -----TOKIQ-----YATEPPE-----QSYLAIPASSW 953
Db 700 EIMRGNQTPGLQGVFSTFSANTPQMLIDIRNKAKALQVDVDEVENTLQSYL-----G 756
Qy 954 VDFIDWLTPSSCCRLYISGPNKDFCPTVNSLN-----CLKNCMSITMGSVRPS 1004
Db 757 VNDP---NFQORTYRVYVQADAQFRSNPEDIGKLYRSANDQMISLSNLVTTITPATGAQT 813
Qy 1005 VEQFHKYLFWFLNDRPNIKCPKGGGLAA--YSTSVNLTSQGVLASRPMAYHKPLNSQDY 1062
Db 814 INHYNLFRSITIN-----GAAAPGYSS-----GQATTAMEKLAKEVLPASMGY 856
Qy 1063 TEALRAARELANITADLRKVPCTDPAF-----EVFPYITITNVFYEQYLTPILPEGLFMLS 1118
Db 857 E-----WSGITAEESGSGQAPLIFGLGLVFPVLVLAQYENYVDP-----IIMLSV 904
Qy 1119 CLVPTFAVSC-----LLGL-----DLRGLNLNLSIVM- 1147
Db 905 FLAILGALSQAQLRGLSNDVFCQVGLVMLIGLAKNAILLIVEFANQLREGGLPIIKATVQ 964
Qy 1148 -----ILVDTVGP-MALWDISYNVSLINLVSAVGSVEPVSHITSFALSTKPTWL 1198
Db 965 ASQERLRPILMTLSLSLTGLW-----PLINPIGA-----GAASRK----- 999
Qy 1199 ERAKEATISGSAVFAVAGVAMTNLPGLVL 1227
Db 1000 -----SLGTAVVGMIVSTVLSLVFV 1020
```

RESULT 52
T20342

hypothetical protein D2013.8 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20342

R:Mortimore, B.

submitted to the EMBL Data Library, January 1995

A:Reference number: Z19259

A:Accession: T20342

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1032 <MIL>

A:Cross-references: UNIPROT:Q18968; UNIPARC:UPI00001642EB; EMBL:Z47808; PIDN:CAA8777.1;

A:Experimental source: clone D2013

C:Genetics:

A:Gene: CESP:D2013.8

A:Map position: 2

A:Introns: 25/2; 73/1; 217/1; 248/3; 570/2; 695/3; 775/1; 807/3; 953/3; 995/3

Query Match 2.0% Score 136.5; DB 2; Length 1032;

Best Local Similarity 19.9%; Pred. No. 0.31;

Matches 100; Conservative 81; Mismatches 199; Indels 123; Gaps 20;

Qy 337 LGQFFQGWGTWASWPLTILVSVIPVVALAAGLVFTETLDPVELWSAPNSQARSEKAF 396

Db 16 IGNAYHDYGRICAHPKACLSMLTMIVLVSPTI-TRLRL-PV---STPIDVFWSEHLH 70

Qy 397 HDQHFGFPFERTNQVILTAPNRSSVRYD---SLLLGPKNFSGILDLDLLLELLEQLERLH 453

Db 71 VNDKIAPFW-INE-----NPASYIQQFIVSTTISPNNATEMGEHAVRAAIAATAFRIQ 123

Qy 454 LQWSPAEQORNISLQDIYAPLNPDTSLYDCCINSL-----LQYFQNNRTL 500

Db 124 ILLAEP-----AVEELCLRLANQRQDSSWPFPSKSLCVLSPASIWNYNLQKFEEDDT 177

Qy 501 LLLTANQTLMGQTSQVDWKHFLYCANAPLTFKDTALALSCMD--YGAPVFPLAIGG 558

Db 178 ITTVFNE-----HCKS---TF-----CMRDLILGAFI---AATG 205

Qy 559 YKGKDYSAEALII---MTFSLNNYPAGDPLRAQAKLWEEAFLEEMRAFORRMAGMFQVTF 615

Db 206 IKQYQTNRKKKIEFAVTMFFARYSKKVIQIREKLQKEFELVDTPPNDQRT--FVQYVF 263

Qy 616 TAERSLEDEINRTTAEDLPATSYIVIFLYISALGSYSWSRWVDSKATLGLGGVAV 675

Db 264 HPLKTFSDYI-----PLIST-YFVCMIVV-----YYSSRKIQMWASRWGLAFASSFT 309

Qy 676 VLGAVMAAMGFYSVLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGPREVH 735

Db 310 VASTLLMTGTICAHLDLSLTTTGWSEVVPYIALIMGLNTLC----- 350

Qy 736 IGRALGRVAPSMLLCS-----LSEALCFELGALTMPAVRTPALTSGL 778

Db 351 ITRSVVYTSPLDVSRRHAHGLSQEGYKLTQYVILELLALLIGLFTRISDIQEFCQFSVI 410

Qy 779 AVILDFLLQMSAFVALLSLDSKR 801

Db 411 CVTVDFYMQLPFYAPCLTFDLQR 433

RESULT 53

G81825

probable drug efflux protein NMA1969 [imported] - *Neisseria meningitidis* (strain Z2491

C:Species: *Neisseria meningitidis*

C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004

C:Accession: G81825

R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel,

; Holroyd, S.; Jørgensen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,

Nature 404, 502-506, 2000

A:Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.

A:Reference number: A81775; MUID:20222556; PMID:10761919

A:Accession: G81825

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1067 <PAR>
 A:CROSS-references: UNIPROT:Q9J7S1; UNIPARC:UPI000000C4CBF; GB:AL162757; GB:AL157959; NID
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: mtrD; NMB1969
 C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 135; DB 2; Length 1067;
 Best Local Similarity 20.9%; Pred. No. 0.41;
 Matches 109; Conservative 83; Mismatches 199; Indels 130; Gaps 25;

QY 452 RHLQVW-SPEAQR--NISLQDICAPLNDPNTSLYDCCINSLLQYFQNNRTLLLTANOT 508
 DB 180 RAMRIWVDPKKLQNTNLSFADVGSA--LSAQNVOISAGSIGSLPAV--RQGT---VTATVT 233
 QY 509 LMGQTSQVMDKDFLYCANAPLTFKDGTTALALSCMADYDAPVFPFLAIGGYKDYSEAR 568
 DB 234 AQQLGTAEFEFGNVILRANT-----DGSNVYLKDVAVRGLGM-----EDYSSST 277
 QY 569 ALI-----MTPSLNNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMPQVTFTAERSLE 622
 DB 278 RLNGVNTTGMVNLN--SGN-AMATAKAVKERMATLEKYFPQGMS--WKTPYDTSKPYE 332
 QY 623 DGINRTAEDLPFATSYIVIFLYISLALGSSYSSSRVMDSKATILGLGVAVLGAVMA 682
 DB 333 ISIEKVIHTLIEMVLVFMVYLFQNI--RYTLPTIVVPISL-----LGGPA----- 379
 QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGR 742
 DB 380 ---FISYMG--SINVLTMPFAMVLVIGIVDDAIVVVENVERIMAGEGLPPKEATKKAMQ 435
 QY 743 VAPSM-----LCSLSEALCPFLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSL 797
 DB 436 ISGAVIGITAVLISVFPVPLAMFSGATGNI--YKQFALTWASSI-----AFSAFLAL--- 484
 QY 798 DSKRQESRLDVCCCKVQDELPPPGQEGE--LLGFFQKAYAPF-----LLHWIT 845
 DB 485 ---TLTPALCATMLKTIPIKGHEEKGGFGFNKFNKNSWTHGYEGRVAKVLRKTP 536
 QY 846 RGVWLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYEFGVAPVYFVT 905
 DB 537 RMVYVIGLAVGV-----FLFMR-----LPTSFLPT 563
 QY 906 LGYNFSSEAGMNAICSSAGC-----NNFSFTQKIQVATEPPE 942
 DB 564 EDQGFV---MVSQVLPAGATQERTNATLAQVTLAKSIPE 600

RESULT 54
 E81051
 multiple transferable resistance system protein MtrD NMB1715 [imported] - Neisseria men
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C:Accession: E81051
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: E81051
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1067 <TET>
 A:CROSS-references: UNIPROT:Q9JY67; UNIPARC:UPI000000C4798; GB:AE002521; GB:AE002098; NID
 A:Experimental source: serogroup B, strain MC58
 C:Genetics:
 A:Gene: NMB1715
 C:Superfamily: acriflavin resistance protein

Query Match 2.0%; Score 135; DB 2; Length 1067;
 Best Local Similarity 20.9%; Pred. No. 0.41;

Matches 109; Conservative 83; Mismatches 199; Indels 130; Gaps 25;

QY 452 RHLQVW-SPEAQR--NISLQDICAPLNDPNTSLYDCCINSLLQYFQNNRTLLLTANOT 508
 DB 180 RAMRIWVDPKKLQNTNLSFADVGSA--LSAQNVOISAGSIGSLPAV--RQGT---VTATVT 233
 QY 509 LMGQTSQVMDKDFLYCANAPLTFKDGTTALALSCMADYDAPVFPFLAIGGYKDYSEAR 568
 DB 234 AQQLGTAEFEFGNVILRANT-----DGSNVYLKDVAVRGLGM-----EDYSSST 277
 QY 569 ALI-----MTPSLNNYPAGDPRLAQAKLWBEAFLEEMRAFORRMAGMPQVTFTAERSLE 622
 DB 278 RLNGVNTTGMVNLN--SGN-AMATAKAVKERMATLEKYFPQGMS--WKTPYDTSKPYE 332
 QY 623 DGINRTAEDLPFATSYIVIFLYISLALGSSYSSSRVMDSKATILGLGVAVLGAVMA 682
 DB 333 ISIEKVIHTLIEMVLVFMVYLFQNI--RYTLPTIVVPISL-----LGGPA----- 379
 QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRPGPREVHIGRALGR 742
 DB 380 ---FISYMG--SINVLTMPFAMVLVIGIVDDAIVVVENVERIMAGEGLPPKEATKKAMQ 435
 QY 743 VAPSM-----LCSLSEALCPFLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSL 797
 DB 436 ISGAVIGITAVLISVFPVPLAMFSGATGNI--YKQFALTWASSI-----AFSAFLAL--- 484
 QY 798 DSKRQESRLDVCCCKVQDELPPPGQEGE--LLGFFQKAYAPF-----LLHWIT 845
 DB 485 ---TLTPALCATMLKTIPIKGHEEKGGFGFNKFNKNSWTHGYEGRVAKVLRKTP 536
 QY 846 RGVWLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYEFGVAPVYFVT 905
 DB 537 RMVYVIGLAVGV-----FLFMR-----LPTSFLPT 563
 QY 906 LGYNFSSEAGMNAICSSAGC-----NNFSFTQKIQVATEPPE 942
 DB 564 EDQGFV---MVSQVLPAGATQERTNATLAQVTLAKSIPE 600

RESULT 55
 H91170
 hypothetical membrane protein [imported] - Escherichia coli (strain O157:H7, substrain :
 C:Species: Escherichia coli
 C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C:Accession: H91170
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
 DNA Res. 8, 11-22, 2001
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
 A:Reference number: A99629; MUID:21156231; PMID:11258796
 A:Accession: H91170
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-749 <HAY>
 A:CROSS-references: UNIPROT:Q8X6L3; UNIPARC:UPI0000165513; GB:BA000007; PIDN:BA837759.1;
 A:Experimental source: strain O157:H7, substrain RIMD 0509952
 C:Genetics:
 A:Gene: EC84336

Query Match 1.9%; Score 134; DB 2; Length 749;
 Best Local Similarity 19.0%; Pred. No. 0.3;
 Matches 162; Conservative 129; Mismatches 313; Indels 248; Gaps 40;

QY 455 QWSPEAQRNLSLQDICAPLNDPNTSLYDCCINSLLQYFQNNRTLL--LTANQTLMOQ 512
 DB 62 QEWLTLQKSAALGSDV-KGPMDDASQAWGA-----FFWQHRNGLDIPNTRARLONG 113
 QY 513 TSQVDMKDFLYCA-----NAPLTKDGTALALSCMADYCAPV----- 550
 DB 114 EAQAWILSQLYSAPSGVSGKELQNDPLMLWRGSQLA---MAKQQRLLMDGWLVTQDP 170
 QY 551 ---FPFLAIGGYKKG--DYSEAEALIMTFSL-----NNYPAGDPRLAQAKLWBEAFLE 599

Db	171	QGNWYLLHGLAGSSDFMQOOTHQIITLNTLEKDLKTRYP-----QAQLLSR-----	218
Qy	600	MRAFORRMAGMFOVTTAERSLEDEINRTTAEDLFIATSYVIVFLYISIALGSSWSR	659
Db	219	-----GTVFSDYASQAQKD-----ISTGLVATLLGVILLIVAVFRSLR	258
Qy	660	VWVDSKATLGGVAVVLGAVMAAMGFPSYIGIRSSLVILQVFLVLSVGADNIPFVL	719
Db	259	PLLLCVISIGICALA---GTVATILLIFGEHLMTLVMSMSVI-----GISAD-----YTL	305
Qy	720	EYQRLPRPGRPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTPMPAVRTFALTSGLA	779
Db	306	YI--LTERWNGNDVSPWQSLAKVNALLLALLTTVAAYLLIMLAPFGIR-----	354
Qy	780	VILDFLLQMSAFVAL--LSLDSKROQASRLDVC-----CCVKPQBELPPPGQEGILLGFFQ	833
Db	355	-----QWAIFAAVGLS-----ASCLTVLFWHPMLCRGLPVRPVPAMA-----	391
Qy	834	KAYAPFLHWHITR-----GVVLLFLALFGVSLYSMCHISVGLDQDELALPKDSYLLDY	886
Db	392	-----LMLRWLAAMRRNKKLSLGLPVALALFSLAGMSMLRVDDDISQLQLPQ-----H	440
Qy	887	FLFLNRYE--VGAPV---YFVTTLIGYNFSSBAGMNAICSSAGCNFSTQKIOVA--TEF	940
Db	441	ILAQEKAITALTGQSVQDKWFV---YGDSPOQTLRRL-----EKYTSLEVAKKEG	489
Qy	941	PRQSYLAIPASSWDDFDIMLTPSSCCRLYISGPNKDKFCPST--VNSLNCCLKNCHSITMG	999
Db	490	LISNRYTIPLNSLARQEDLQL-----LKTAAPTPTKALQNAGLTAVNPDLNAMPVND	543
Qy	1000	S--VRPSVEQFKHYLPWFL--NDRPNIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKP	1055
Db	544	EWLASPASSEGMR--LLMLTLNENGESGVLPVVEG-----VKSSALM-----	581
Qy	1056	LKNSQDYTEALRAARELANIATADLRKVPGT---DPAPFVFPYTIINVFYEQYLITLPEG	1112
Db	582	-----QEIATYPCGTAWVDKSTFPDELALRYVLT-----G	614
Qy	1113	LFMLSLCVPTFAVSCLLGLDLRGLAMLSIVMLIVDTVGFMAIMDISYNAVSLINLV	1172
Db	615	LLVALAVIACGAVA---RLGWKGLISLPSVLSLGCGLAVLAMSGQAVNLFSLALV	670
Qy	1173	SAVMSVFEV-----SHITRSPAISTKPTWLERAKEATISMGSAVPAGV--AMTNLPG	1223
Db	671	LVLGIINTLFPSPNPGTPLTSLALAL-----AMLTTLTLGLMLVFSATQALSSFGI	724
Qy	1224	ILVLGLAKAQLI	1235
Db	725	VLVSGIFTAFLL	736
RESULT 56			
D69403			
conserved hypothetical protein AF1229 - Archaeoglobus fulgidus			
C;Species: Archaeoglobus fulgidus			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004			
C;Accession: D69403			
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson			
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.			
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.			
Nature 390, 364-370, 1997			
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.			
Smith, H.O.; Woese, C.R.; Venter, J.C.			
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae			
A;Reference number: A69250; MUID:98049343; PMID:9389475			
A;Accession: D69403			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-750 <L>			
A;Cross-references: UNIPROT:O29039; UNIPARC:UPI0000056DE1; GB:AE001019; GB:AE000782; NID			
C;Superfamily: Archaeoglobus conserved hypothetical protein AF0459			
Query Match 1.9%; Score 133; DB 1; Length 750;			

Best Local Similarity 19.5%; Pred. No. 0.36;			
Matches 137; Conservative 117; Mismatches 273; Indels 176; Gaps 33;			
Qy	597	LEMRAPORMAGMFOVTTAERSLEDEINRTTAEDLFIATSYVIVFLYISIALGSSYSS	656
Db	166	VERKISFTNPPHGLV--LQLTGGPALSYIEREIKSGFI--TMMVSIILMLVLLFLTFSG	222
Qy	657	WSRVVDSKATLGGVAVVLGAVMAAMGFPSYIGIRSSLVILQVVPFLV---LSVGADN	713
Db	223	-----AVRKKITAFPLVISVM--SVTVVGLMPILGIPLSEHTNGALPMLGLAIEYGAQ-	276
Qy	714	IFIEVLEYQRLPRPGRPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTP-MPAVRTF	772
Db	277	---LQNRYYE--ERREGRDVDAVVISITRTGLAIVM-ALITTVIGFMSMLAPGMPAMAQF	331
Qy	773	ALTSGLAVIDFLQMSAFVALLSL-----DSKROEASRLDVCCCKVPQELPPPGQEG	827
Db	332	GIISLGLIVAYLLTTLTFLPAVLKLDHWNDAKEQK-----REBEK	372
Qy	828	LIQFFQKAYAPFLHWHITRGVLLL---FLALFGVSLYSMCHISVGLDQDELALPKDSYLL	884
Db	373	SVGTLENSLSAISITLRTATRPVGLVAASIIVLFG--LYAAPQIGLETNYNKKYVPQDLTAM	430
Qy	885	DYFLFLNRYEFGVAPYFVTTTLGYNFSSBAGMNAICSSAGCNFSTQKIQATEFFPQS	944
Db	431	QRFKEIERL--VGGQATYTLVLEVD-----ELNA-----ETLKEIDELS	467
Qy	945	YLAIPASSWDDFDIMLTPSSCCRL-----YISGPNKDKFCPSTVNSLNCCLKNCS	995
Db	468	KYIVKEBELIYD-----SSITKLISEVRKAVGFEGLPESDAELMQILSAL-----	513
Qy	996	ITWGSVRPSVEQFKHYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKP	1055
Db	514	-----PQEVSRVI-----SGGQIAVHFSSNADSQDE-----YISTHK-	546
Qy	1056	LKNSQDYTEALRAARELANIATADLRKVPGTDPAFVFP-----PYTIIN--VFYEQYLITL	1109
Db	547	-----SIVRDV-----EFFGMSGYYVTGPPVIYGEIMGRLM	577
Qy	1110	PEGFLMLSLCVPTFAVSCLLGL--DLRGLMLLSTVMIL--VDTVGFV-----ALWD	1160
Db	578	TSG--QTTMTVAVALLIMLLAVYRSIRKAVPFLIATSVIGAMNTIMFLTGYKQTWIS	635
Qy	1161	ISYNAVSLINLVAVMSVFEVSHI--TRSPAISTKPTWLERAKEATISMGSAVFAGVANT	1219
Db	636	IAMSITL-----GLGIDFSIHVLERYFEERAAPSPVEAVRRTIERTKAITTS--ALT	687
Qy	1220	NLPGLVLGLAKAQLIQIFPFRMLNLLITLLGLLHGLVFLPVIL	1262
Db	688	MAGFGSLMFTFPIQNFGE-IALVALFLSLAALTVPVPAPL	729
RESULT 57			
C69307			
conserved hypothetical protein AF0459 - Archaeoglobus fulgidus			
C;Species: Archaeoglobus fulgidus			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004			
C;Accession: C69307			
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson			
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.			
.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.			
Nature 390, 364-370, 1997			
A;Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.			
Smith, H.O.; Woese, C.R.; Venter, J.C.			
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae			
A;Reference number: A69250; MUID:98049343; PMID:9389475			
A;Accession: C69307			
A;Status: preliminary; nucleic acid sequence not shown; translation not shown			
A;Molecule type: DNA			
A;Residues: 1-736 <L>			
A;Cross-references: UNIPROT:O29790; UNIPARC:UPI00000570DE; GB:AE001073; GB:AE000782; NID			
C;Superfamily: Archaeoglobus conserved hypothetical protein AF0459			

Query Match 1.9%; Score 132; DB 1; Length 736;
Best Local Similarity 19.3%; Pred. No. 0.41;
Matches 150; Conservative 107; Mismatches 314; Indels 208; Gaps 28;

```
QY 558 GYKGDYSEAEALITWTSGLNNYPAGD-----PRLAQAKLWEAF----- 596
DB 105 GRDGVSTSSPASIRDVLDGDIPOSEELVGTABRLAKSLLPERSFALISVQIESMENDR 164
QY 597 -----LEEMRAFQRRMAGNFQVTTAERSLEDEINRTTAEDTLPFATSPYIVIFLYISL 649
DB 165 LWELAEERIERILQFVSPPGVV--VDVTCGAVLWYQIAKAVEED--VYRTFFIASVTMI 221
QY 650 ALGYSYSSWRVNDKATLGGVAVVLGAVMAAGPFSYIGIRSLVILQVVPFLVLSV 709
DB 222 LVAVFSG----VVRSKLTV-LFPLIVSLMSVAVMAGLLPLIGLQMTSYNSATIPILI-GL 275
QY 710 GADNIFIVLEVQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAV 769
DB 276 AIDVAAQLONRYER-ERGGREVDALFEELNHTRFFLFMAMATTVIGFASMCAPGIPSL 334
QY 770 RTPALTSGLAVIDFLQMSAFVALLSDSKRQEASRLDVCCKVKPQELPPPGGEGILL 829
DB 335 FWFPSLLSVGLITAFLLSNTPLPSSLKLSOR-----GGREESISP 374
QY 830 GFPOKAYAPFLLHWITGV-----VLLFLALPGVSLYSCHVISGVLDOELALPKDSYL 883
DB 375 GILEK-----LLTTIAKVTIANSRKILALTILILVVGAVASTHVELETRNRKVPAPDL 429
QY 884 LDYFLPLNRYEVCAP--VFVTTLGYNFSEAGMNA-----ICSSAGCNGNPSFTQIKI 935
DB 430 LVKEEERKV---SPQVVALVLSRTERIDAESVVKAEELAKYISSREIYDYSLGKTL 486
QY 936 YATFPBQSYLAIPASSWVDVDFIDWLPSSCCRLIYISGPNKDKPCPSTVNSLCLNCKMS 995
DB 487 -----EYVYGAIPAESF-SKAIDAL-PKSIYSRPFISGNQLAVFYVNTQTREEVEK 538
QY 996 ITMGSVRPSVEQFHYLPWFILNDRPNTKCPKGLAAVSTVNLTSQGVILASRFMAYHKP 1055
DB 539 -----SIENDLRPLEW-----DGD----- 552
QY 1056 LKNSODYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFEQVLTILPEGLF- 1114
DB 553 -----YYITG--FPVILGHLSEVLFS 571
QY 1115 MLSLCLVPTFAVSCLLGLDLRS---GLNLLSTVMILVDVTGFMALWDISYNAVSLNL 1171
DB 572 SLNLTMLVAILLIVTFLLAYRSVVRAIILPISISVAVSMNVCMLLMGIKQTTSIALN 631
QY 1172 VSAVGMSEFVSHITRSFAISTKPTWLERAKEATIS-MGSAVFAGVAMTNLPGLVLGLA 1230
DB 632 SILVGMGIDYSIHITEREYPERGKFGVEESVKRTVERTGKAVLTS-ALTAGGFGAL--- 687
QY 1231 KAQLIQIOPFPLNLLITLLGLLHGLVFLPVLISVVGPDVNPALAEOKRABEAVAVMV 1289
DB 688 -----YFSTPVLNSFNGI--LAFIAIFSLLA-----ATSAPVALLI 722
```

RESULT 58
C96945
cation efflux system protein, AcrB/AcrF/AcrD family [imported] - Clostridium acetobutyli
C/Species: Clostridium acetobutylicum
C/Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C/Accession: C96945
R/Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A/Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A/Reference number: A96900; MUID:21359325; PMID:21359325
A/Accession: C96945
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-1166 <KUR>
A/Cross-references: UNIPROT:Q97M30; UNIPARC:UPI00000C9E4A; GB:AE001437; PTDN:AAK78350.1


```
QY 1117 SLCLVTPFAVSCLLGL-----DLRSLNLLLSIVMLIVDTVGFMAWLDISYNAVSLI 1169
DB 846 -SALYATFGLALAIIFLVLAIOFESLRDPLVIMVSVPLAICGALIALAWGTATMNIYSQV 904
QY 1170 NLVSAVGMVSE---FVSHITRSPAISTKPTWLBRAKEAT-----ISMGSVAFAGVAMT 1219
DB 905 GLITVGLITKHCILICEVAKERQHNKLSRIEAVNHAARVLRPRIMTMTAMIAGL----- 961
QY 1220 NLPGLIVLGLAKAQLIQIPFRLMLLIT---LGLLHGLVFLPVLISYVGPVNP-ALAL 1275
DB 962 -IPLMAYATGAGAOR---FSIGIVIVSGLAIGTFTFLVLPVIYSYLAERKHKLPLPVFV 1015
QY 1276 EQK-----RAEBAVAA 1286
DB 1016 EDKLEKLARIDEAKAA 1032

RESULT 61
H86016
hypothetical protein Z4861 [imported] - Escherichia coli (strain O157:H7, substrain BDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: H86016
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: H86016
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-772 <STO>
A:Cross-references: UNIPROT:Q8X6L3; UNIPARC:UPI0000165968; GB:AE005174; NID:g12518149; B
A:Experimental source: strain O157:H7, substrain BDL933
C:Genetics:
A:Gene: Z4861

Query Match 1.9%; Score 130; DB 2; Length 772;
Best Local Similarity 18.8%; Pred. No. 0.6;
Matches 160; Conservative 128; Mismatches 317; Indels 246; Gaps 39;

QY 455 QWSPBAQRNISLQDIQYAPLAPNDTSLYDCCINSLLQYFQNNRTLL--LTANQTLMQQ 512
DB 85 QEWLTLLQKSAALGDV-KGPMDAASQQAQWGA-----FFWQHRNGLIDPNTFRLQNGG 136
QY 513 TSQVDMDKDFLYCA-----NAPLTFKDGTLALSCMADYGAPV----- 550
DB 137 EAQAQWILSQAFAFSGVSGKELQNDPLMLMRGSQA---MAKNGQRLMDGWLVTQDP 193
QY 551 ---FPFLAIGGYKKG--DYSEAEALIMTFSL-----NNYPAGDPRLAQAKLWEEAFLEE 599
DB 194 QGNYWYLLHGLAGSSFDQOHTLITLTKLEKOLKTRYP-----QAQLLSR----- 241
QY 600 MRAFORMAGMQVTTAARSLEDEINRTTABDLPIFATSYIVIFVYISLALGSYSSWR 659
DB 242 -----GTVFYSDYASQAQKD-----ISTLGVATLIGVILLIVAVPRSLR 281
QY 660 VMVDSKATGLGCVAVVLGAVNMAAGCFPSYLGIRSLVILQVVPFLVLSVGADNIFIFVL 719
DB 282 PLLLCVISIGIGALA---GTVATLIFGELHMLTLVMSMSVI-----GISAD-----YTL 328
QY 720 EYQRLPRRPGEPREHVHIGALGRVAPSMLLCSLSEACFFLGAITPMPAVRTFALTSLGA 779
DB 329 YY--LTERVHVGNDSVPWQSLAKVRNALLALLTVAAYLIMMLAPPFGIR----- 377
QY 780 VILDFLLQWSAFVALLSLDSKQKQEARLDVC-----CCVKPQELPPPGQEGILLGFQK 834
DB 378 -----QMAIFAAV-----GLXASCLTVLFWHPWLCRGLPVRPVPAMA----- 414
QY 835 AYAPFLHWTIR-----GVVILLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYF 887
DB 415 ----LMLRWLAWRNKKLSLGLPVALALFSLAGMSMLRVDLDDISQLALPQ-----HI 464
```

```
QY 888 LFLNRYFE--VGAPV---YFVTTLGVNFSSEAGMNAICSSAGCNNSFTQKIQYA-TEFP 941
DB 465 LAQEKAITALTQSDVDQKWFV---YGDSPQOTLRRL-----EKYTASLEYAKKEGL 513
QY 942 EQSYLAIPASSWDDPIDMLTPSSCCRLYISGPNKDKFCEST-VNSLNCNKNCSITMGS 1000
DB 514 ISNYRTIPLNSLARQBEDLQL-----LKTAAPTVTTKALQNAAGLTAVNPDNLAMPVNVDE 567
QY 1001 ---VRPSVEGFHKYLPWFL--NDRPNIKCPKGGLAAYSTSVNLTSDQVLASRFMAVHKPL 1056
DB 568 WLASPASEGWR--LLWLTLENGESGVLPVEG-----VKSSALM----- 604
QY 1057 KNSQDYTEALRAARELANITADLRKVPGT--DPAFVFPYTTITNVFEQYLTILPEGL 1113
DB 605 -----QEIATYPQCGIAWDRKSTFDELFALYRYVLT-----GL 638
QY 1114 FMLSCLVPTPEAVSCLLGLDLRSGLNLLSVMLIVDTVGFMAWLDISYNAVSLINLVS 1173
DB 639 LLVALAXIACGAVA---RLGWRKGLISLPSVLSLGCGLAVLAMSQQAVNLSLALVL 694
QY 1174 AVGMSVEFV-----SHITRSFAISTKPTWLBRAKEATISMGSAVFAGV-AMTNLPGI 1224
DB 695 VLIGIGINYLPFSNPRGTPLTSLALAL-----AMLTTLLTGLMLVFSATQAISSFGIV 748
QY 1225 LVLGLAKAQLI 1235
DB 749 LVSGIFTAPLL 759

RESULT 62
A89862
Na+/H+ antiporter subunit [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: A89862
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogura,
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89862
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-801 <KUR>
A:Cross-references: UNIPROT:P60674; UNIPARC:UPI00000CAC59; GB:BA000018;
A:Experimental source: strain N315
C:Genetics:
A:Gene: mnhA
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 1.9%; Score 129; DB 2; Length 801;
Best Local Similarity 18.7%; Pred. No. 0.74;
Matches 131; Conservative 99; Mismatches 202; Indels 268; Gaps 31;

QY 641 IVIFVYISLALGSYSWSRVM-----VDSKATLGLGCVAVLGVNMAAGCFPSYLG 692
DB 39 IVIFVYMLTLIKTYTSGNTVMKTLNWMHPFGNFDLYDLGLGLSLISIGTG----- 91
QY 693 RSSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREHVHIGALGRVAPSMLLCSL 752
DB 92 --SLVLYSIGYLSKSEQLGNFYCYLL----- 116
QY 753 SEACIFFLGAITPMPAVRTFALTSLGLAVIL-DPFLQMSAFVALLSLDSKQKQEARLDVCC 811
DB 117 -----LPMGAM-----LGWVLSNDNIVILYLFWELTSPSS----- 145
QY 812 CVKPELPPPGQEGILLGFQKQKAVAPFLHWTITGVWLLLF--LALFGVLSYMSCHISV 869
DB 146 -----FLLISFWERQAS--IYGAQSLIITVFGSLSLGGLIIL----- 182
QY 870 GLDQELALPKDSYLLDYFLFLNRYFEVGPVYFVTTLGVNFSSEAGMNAICSSAGCNNS 929
```

Db 183 -----LAIPQSFQIY--MIQHASEIQNSPFFIFAM-----ILIMIG-----A 219

Qy 930 FTQKIYATEPEOSYLAIPASSWDDFDIDWLTSSC-----CRLYISGNKKKFC 980

Db 220 FTKSAQF-----PFYIWLDPAMEAPTFSAYLHSGATWVKAGLYLIARMTIFA 267

Qy 981 PS-----TVNSINCLKNCMSITMGSVRPSVQFHKYLPWFILNDRPNIKCPKGGGLAAVSTS 1035

Db 268 ASQGWVTVT-----LVGLITLFWASLNATKQDDUK-----GILAFSTV 306

Qy 1036 VNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAANI-----1076

Db 307 SOLGIMAMWLGISYHYQGDDSKIYAAFTAAIFHLINHATPKGALFMITGAVDHSSTG 366

Qy 1077 TADLRKVPGTDPAFVFTYTNVYEQYLILPEGLFMLSCLVPTFANVSCLLGLDLR 1136

Db 367 TRDVKKLGG-----LITIMPISFT-----ALSMAGVPPF-----400

Qy 1137 SGLNLLSTV-----MLVDTVGFMALWDISYNAVSLINLVSAGVMSVEFVSHIT 1186

Db 401 NGFLSKESLETFTTASQANLFSVDTLGIL-----FLIIGVSVFTFYVSIKFIHIL- 453

Qy 1187 RSFAISTKTWML-ERAKEATISM--GSVAFAG--VAMTNLPGILVLGLAKAQLIQI-----1237

Db 454 --FFQYKPEQLPKKAHEVSIILMLSPAILATLIVILGIFPGILTNISIIEPATSSINHTV 511

Qy 1238 -----PF-----FRNLTLITLGLLGLVFLPVILSY 1264

Db 512 IDDVBFHMFHGLTPAFLSTLVIYILGIL-----LIVTFYS 546

RESULT 63

AB0370

multidrug efflux protein [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C:Accession: AB0370

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titchell, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001, MUID:21470413, PMID:11586360

A:Accession: AB0370

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1043 <KUR>

A:Cross-references: UNIPROT:Q8ZCB8; UNIPARC:UPI00000DCAD6; GB:AL590842; PIDN:CAC92285.1;

C:Genetics:

A:Gene: YPO3043

C:Superfamily: acriflavin resistance protein

Query Match 1.9%; Score 129; DB 2; Length 1043;

Best Local Similarity 18.8%; Pred. NO. 1.1;

Matches 185; Conservative 146; Mismatches 287; Indels 366; Gaps 54;

Qy 475 LNPNTSLVDCGNSLLQFQNNRTLLLL-----TANQTLMGQTSQVDWKDHF 522

Db 188 LDPNKLNTYQLTSDIVSAIQSQNTQVAVGQLGGTTPAVDQNALNATINAQ-SOLOTPPEF 246

Qy 523 LYCANAPLTFK---DGTALALSCMA--DYGAPVFPEL-----AIGGYKGDYSEAEALI 571

Db 247 -----REITLRVNQDSGLVTLGDVAKIEGSEKYDLSRFNGQAASGMGKILASGANEL- 300

Qy 572 MTFSLNYPAGDPRLAQAKLWEAPLEEMRAF-----QRRMAGMFOVTTAERSLEDEIN 626

Db 301 -----QTDKRV-----KARLAELAPFPFHGLEAKIA--YETTPFVQASIKDVV- 341

Qy 627 RTTAEDLPFATYSYIVIFLYISALGYSWSRWDVKATLGLGVA---VVLG--AVM 681

Db 342 KTLLE-----AILVFLVWYLLQNF-----RATL-IPTVAVPVVLLGTFVAVL 383

Qy 682 AAMGFFSYLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQLRPRPG-EPREV---HIG 737

RESULT 64

DB3393

RND multidrug efflux transporter PA2018 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C:Accession: DB3393

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B. adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, ; Lory, S.; Olson, M.V.

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho Nature 406, 959-964, 2000

A:Reference number: AB2950; MUID:20437337; PMID:10984043

A:Accession: DB3393

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1045 <STO>

A:Cross-references: UNIPROT:Q9RG59; UNIPARC:UPI00000D420B; GB:AE004628; GB:AE004091; NID

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2018

C:Superfamily: acriflavin resistance protein

Db 998 HAVGTGVGMGMSATVLAIPFVPL 1021

Qy 1228 -----GLAKAQLIQIFPRL 1242

Db 998 HAVGTGVGMGMSATVLAIPFVPL 1021

555 EIEHEPGTSLERTRETVEALIEQKLEDYSEIENFVSTSGARGTNNVRRDNGNSYEGQVQATL 614

904 -----TTLGYNFSSEAGNNAICSSAGCN-----NFSFTQKIOVA 937

615 VPSDDRDLTLTLEPVESIKRDVERVDSDAEINVSPLSLQSGSEPNTLTFYLSDSDDPDLSEA 674

938 TEFPQSYLAIPASVWDDFIDMLTPSSCCRLYISGPNKDFKPCSTVNSLNCLKNCMGIT 997

675 TEALBEELLERNEVTNLELSIDVTTPEYQLLIDEEAARENGLVPAQVAF-----VYDVT 729

998 MGSVRSVQFHKYLPWFANDRNKICPKGGLAAYSTSVNLT-----1039

730 LQGTVTLEIDNDLYDLFVYHFDV-----LDSLETWEELTIRNNEGYYVPLGDVVITIE 783

1040 -SDGQVLASRF-----MAYHKPL-----KNSQDYTE-ALRAARELA 1073

784 EGECPATINRIEQDRAVEFTVYDSSLSLGDISSIVLDIVDEQNFEDTTEFAFTGDQAL 843

1074 ANITADLRKVPGTDPAPFVPPYITNVFYQY-----LTILPEGLFMLSCLVPTFAVS 1127

844 ETAMADL---TCGAFMLAVFVFLVMAAQFSEFRFPVFMFSLPLVMIGVSLALLVTGTSF 900

1128 CILLGLDRLSGLNLSIV---MILVDTVG-FMALWDISYNAVSLINLVSAGMGSVEFV 1182

901 SVMAGI-----GMILAGIVVNNNAIVVDYNNQLKASGATSYEAI-----VEAV 944

1183 SHITRSFAISTKPTWLERAKEATISMGSAFVAGVAMTNLPGILVLGLAKAQLIQIFPFL 1242

945 KQTRPVLMTALTIL-----GVLPLALGEGTEMORPIGLTVIG-----985

1243 NULLITLLGLLHG-----LVPLPVTLS-----YVGPD--VNPALALEKRAERA 1283

986 -----GLISGFTFLFVIPVIYSLFDPKQTRLNKVTYITPDGELLIPARLLQERKOLEA 1037

RESULT 68

G87236

C:Species: Mycobacterium leprae

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C:Accession: G87236

R: Cole, S.T.; Eiglmier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Haim, M.A.; Rutherford, K.M.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, Nature 409, 1007-1011, 2001

A: Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sg, A: Title: Massive gene decay in the leprosy bacillus.

A: Reference number: A86909; PMID: 21128732; PMID: 112334002

A: Accession: G87236

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-1013 <STO>

A: Cross-references: UNIPROT: O06079; UNIPARC: UPI000016SF6D; GB: AL450380; NID: g13093843; F: C: Genetics:

A: Gene: ML2617

Query Match 1.8%; Score 126.5; DB 2; Length 1013;

Best Local Similarity 18.9%; Pred. No. 1.5;

Matches 138; Conservative 99; Mismatches 274; Indels 221; Gaps 29;

Qy 632 DLIPATSYIVIFLYTISLALGSYSSWRVMSVDKATLGLGCVAVVLG--AYMAAGFFPSY 689

Db 184 NLPITILVLAVF-----GSLAAAVPLALGVCTVVVTMGLVDL 222

Qy 690 LGIRS--SLVILQVPLVLSVGADNIPFVLEYQRLPRRGPBPREVHIGALGRVAPSM 747

Db 223 VSMHTMSVFTSTVSMFGIALAVDYSFLIMRPRELRSRGPQEA-VDAAMATSGLAV 281

Qy 748 LKCSLSEACFFLGALTPPAVTEPALT--GLAVILDFLLQMSAPVALLSDSKREQASR 806

Db 282 VLUGMT-----VVASUTGYLINTAAKSWATGAILAVATMLASITL-----324

Qy 807 LDVCCCVKQELPPPPQCG---EGLLGLFFQKAVAPFLLHMTITRGVLLFLLFALFCVSLYS 863

Db	325	-----TPAALATFGRAAAVKRSVLMHWSORSECTQSLFW--TRVWGVWHRPWSASAAS	376
Qy	864	MCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTGLGYNPSE--AGNNAICS	921
Db	377	TLILII-----MATPVTSMGLNSLLRQ-----FNSSHETRAGVAAARQ	414
Qy	922	SAGCNFNSTFKIQIYATEPPEQSYLAIPASSWVDDFDIMLTPSSCCRLYIIGPNKDKFCP	981
Db	415	ALGPGCALG-----PVQVLITPP-----DD-----PNTQASSP	441
Qy	982	STVNSLNCILKNCM--SITMGSVPSPVEQPHKYLFPWFLNDRBNIKCPKGLAAYSISVNLIT	1039
Db	442	KHRQTIGAIRNRMLQAKNWSVAP-----PQPADN--NCSALLGAV-----LS	482
Qy	1040	SDGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPEVFPYTTIN	1099
Db	483	VDPEDL-----GAREYDVMRTELPKVPGAHVNVGGTALIN	520
Qy	1100	VPEYQVLTLLPEGLFMLSICLVTFPAVSCULLGLDLRSGLNLLSI-----VMIL----	1149
Db	521	DPDRVAKTEPLMLVFLVALLAFVMLATISIRSVFLAPKGLVMTLLSVAAAAYGSLVMVFWG	580
Qy	1150	-VDVTGFMALWDISYNAVSLINLVSAGMSVEP-VSHITR-----SPAIS	1192
Db	581	WLENIGFTHINSIDSTVPPPLV-LAMTFGLSMDYBEIFLLTRIRERFLOTGHTRDAVAGVS	639
Qy	1193	TKPTWLERAKEATISMGSAVFAVGMVMTNPLGILVLGLAKAQLI--QIPFFRLNLLITLIG	1250
Db	640	TSARIITSAALIMI-----AVFGFAPAGMLPVAHIGVACAVAIADVTAVRLVLPVTLMA	695
Qy	1251	L-----LHGLVFLPVLISYVG---PDVNPALALEOK	1278
Db	696	MFAQNMWMLPRLMSRALPAVDKPPPPVDLNEIVLVPADISATKVPQCDLRMVLKLAAK	755
Qy	1279	RAEEAAVAAMVA	1290
Db	756	LKNLAPDAICVA	767
RESULT 69			
D72329			
hypoetical protein - Thermotoga maritima (strain MSB8)			
C:Species: Thermotoga maritima			
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004			
C:Accession: D72329			
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke-			
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D. C.M.			
Nature 399, 323-329, 1999			
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome se-			
A:Reference number: A72200; MUID:99287316; PMID:10360571			
A:Accession: D72329			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-674 <ARN>			
A:Cross-references: UNIPROT:Q9WZ84; UNIPARC:UPI00000C137D; GB:AE001749; GB:AE000512; NTI:			
A:Experimental source: strain MSB8			
C:Genetics:			
A:Gene: TM0817			
Query Match			
Best Local Similarity 18.0%; Score 126; DB 2; Length 674;			
Matches 130; Conservative 112; Mismatches 259; Indels 220; Gaps 29;			
Qy	558	GYKGD-KYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLTEEMRAFORMAGMFQVTF	616
Db	152	GVNARDKPIEKRLVSSYEKNYLFGEPI-VIDSTLKE-----LVKQIFV	195
Qy	617	AERSLEDEINRTTAEDLPATSYIVIFLYISIALGYSYSSWRVMSKATLGLGGVAVV	676
Db	196	-----YPVF--MFLVIFLLFYTLQRSFRA-----AIFSL--IVPV	226

```
QY 677 LGAVMAAGPFSSYLGISSVLQVWPELVLSVADNIFIFVLEQRLPRPBGPREVHI 736
Db 227 LATFFVFAVFA-MCKSLNTVMTITPELLI-IGS-----AYGLHYNALFRFSKRE--- 277
QY 737 GRALGRVAPSMLLCSLSBAICFFELGALTPMPAVRTEAL--TSGLAVIDLDFLLQMSAFVAL 794
Db 278 --AVKHKPKILFMSLTTAAGFMSVFDIDIRAFRELGLVSGSLAVVVLVIF----- 327
QY 795 LSLDSKQEARLDVCCCKFQELPPPCQGGGLLGFQKAYAPPELLHWITRG--VVILL 851
Db 328 -----TSGVEIFRNYTPKRT-----RSFGMKYVGRKIALIVLV 361
QY 852 LFLALFGLVSLYSMCHISVGLQDELALPKDSYLLDYFLFLNRYFVEGAPVYFVTTLGNFS 911
Db 362 VFLVMAALSPFLKRVQVGSNDVSYFERDSERKAYDLIVKFKNTREPIYLVLKGNPFV 421
QY 912 SEAGMNAICSSAGCNCNFSFTQKIQYATEPPEQSVLAIPASSWVDDFDIMLTPSSCCRLYI 971
Db 422 G-----TDSKILKELI 432
QY 972 SGPNKDKFCPSTVNSLNCNKMSITMGSVRPSVEQFHKYLPWFNLDRPNIKCPKGLAA 1031
Db 433 EKIEKSEYSSVVFVDIPVPIIM-YTLGRTPN-----FLKTFVGRNRIR----- 476
QY 1032 YSTSNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL---RKVPGTRDP 1088
Db 477 --LIVNLTPGY-----EHVKVVDLINEVSVSTGSHYVAGS-- 512
QY 1089 APEVFPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLSIVMI 1148
Db 513 -----VLINWDINESIMRSQISQSVLASILI---FAMVFII---FRRLPTLTSVMPI 559
QY 1149 LVDVTG---FMALWDISYNNVSLINLVSAVMSVFEVSHITRSPAISTKPTWLERAKEAT 1205
Db 560 APTTFVNFELFMALFGISLDVSTSIISGLMGLVIDYSIHASEERLRDPLVWK----- 614
QY 1206 ISMGSAVPAGVAMTNLPGLVLGLAKAQIQLIFFFLNLLITLLGLLHGLVFL---PVI 1261
Db 615 -NVGFSV-----LINALG-LISGFVALLFSELALFNISLMLGIGVGAFTLIVQPMI 667
QY 1262 L 1262
Db 668 L 668
RESULT 70
G87398
AcB/AcrD/AcrF family protein [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87398
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; PMID:21173698; PMID:11259647
A:Accession: G87398
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1047 <STO>
A:Cross-references: UNIPROT:Q9A8Z1; UNIPARC:UPI00000C72B9; GB:AE005673; MID:g13422529; E
C:Genetics:
A:Gene: CC1204
```

Query Match 1.0%; Score 126; DB 2; Length 1047;
Best Local Similarity 19.7%; Pred. No. 1.7;
Matches 198; Conservative 138; Mismatches 343; Indels 326; Gaps 50;

QY 500 LLLLTANQTLMGQTSQVDWKDFLYCANAPLTFKDGTFALALSCMADYAGAPVPPF----- 553
Db 137 IIIILNLTSTTLNPLQADYADRYLV---ERMSTVDGVAQANL-----FGAKIYAMRIWLDA 189

```
QY 554 --LAIGYKGYKQYSEAEALIMTFSLN-----NYPAGD-----PRLAQA----- 589
Db 190 NEMAARGVTVDVESA-----LNAQNELPAGSLESAAKDFTIRVARSYATPBEFAR 241
QY 590 -----KLWEEAPLEKRAFQRM---AGHFQVTF-----TAER 619
Db 242 LPLRAANSQSVLRLDGVARIEEGPDPERRRLFRNGIDQVIGLITROSQANDVAISKAVR 301
QY 620 SLDEINRT-----TAEDL-PIFATSYIVIFLYISIALGYSYSWSVMV 662
Db 302 AEVEALNQTLPGTKMVIADNSVPTAEAIHEWITMGLAIGLVVMVNLPLGWSRSALI 361
QY 663 DSKATILGLGGVAVVLGA--VMAAMGFSSYLGIRSSLVILQVVPFLVLSVG--ADNIIFV 718
Db 362 PS-----IVAPICILSTFIILAPLG-----SLNLTLLA-LVLAVGLWDDAIVVV 407
QY 719 LEVQLRR--PGEPEVHIGRALGR-----VAPSMLLCSLSEAI CFFL GALTMPAVRT 771
Db 408 ---ENIQRRVDDGEPPLVAAERGTQVFAVAVTTVLVSVFAPLMPFGTIGRLFVELA 464
QY 772 FALTSGLAVIDLFLQMSAFVALLSLDSKRQEARLDVCCVKPQELPPQGGSEGLL--- 828
Db 465 VAIAAV-----AFSALLALSLSPMWASK-----LLRPAHGGSLARR 502
QY 829 ----LGFQKAYAPPLH-----W-ITRGVVLFLALFGVLSYMSCHISVGLDQELALPK 879
Db 503 VNKTMGALSASRASLEBQILGGWSAAAGVLLVLLAGFAAVLF-----VSLPKELVPAE 556
QY 880 DSYLLDYFL-----FLNRYFEVGAPVYFVTTL-----GYNFSEAG 915
Db 557 DRGRVDVSI SAAEGAGYDYTSAITAKTEKLDKTKREAGVSERTIMTIPRFGNSFNANG 616
QY 916 MNA-----ICSSAGCNCNFSFTQ-KIQYATEPPEQ-----SYLAI 948
Db 617 VVALKPMGERDKTAEVAEALNKELSRFTSVRAVASVRGPFQRCGGGGGGGTNVDFIAT 676
QY 949 PASSWVDDFI---DWLTP-----SSCRLVISGPNKDKFCPSTVNS 986
Db 677 G-----NDYVQLANWLKPIILAEADNPGLARPRLDYEPTAPRLVQ-IDLDKAAATLGV-S 729
QY 987 LNCILKNCMSITMGSVR-----PSVEQFHKYLPWFNL-----DRNLIKCPKGLAAYS 1033
Db 730 AQSVGRALETWFGSRRATYIKSQEYDVLQTNLDQRRSIEDLNRNVRTNTGAI VPLS 789
QY 1034 TSNVLTSDQVQLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAPFVF 1093
Db 790 TVVTELRG-----DTPDRPRVDRLSVTLTAQLN--FG----- 821
QY 1094 PYTTN--VFYEQVLTILP-----EGLFMLSCLVPTFAVSCILLGLDL 1135
Db 822 -YTVEDAVAFPKDQAACHPAPGVSVKWGGQAKDYLEGSGGIAI----AFGLALLLVFLVL 876
QY 1136 RSGLLNLL--SIVMLVDVTGFMALWDI-----SYNAVSLINLVSAVGMS-----VEF 1181
Db 877 AAQFESWIHPAVIMLTVPLAALGUGLMTGSTINTYSGIILIGIATKNGILIVVEP 936
QY 1182 VSHITRSPAISTKPTWLER--KEATISMGSAVPAGVAMTNLPGLI--VLGLAKAQILOI 1237
Db 937 ANQL-RDEGLSVREAVIEAALRLRPIIMTS---VSAAMGALPLMWAGAGAGSRQTICA 992
QY 1238 FFFRNLNLTLLGLLHGLVFLPVILSVYGPDPVNPALALEQKRAEE 1282
Db 993 VIPTGAIPATLTLTFIVPVFDLLARFTKSPETWARQIEAVEAQE 1037
```

RESULT 71

A75560
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: A75560
R:Whitce, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma

S.: Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A>Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: A75560
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-749 <WHI>
A:Cross-references: UNIPROT:Q9RY43; UNIPARC:UIP00000C16CE; GB:AE001873; GB:AE000513; NID:
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0108
A:Map position: 1

Query Match	1.8%;	Score 125.5;	DB 2;	Length 749;
Best Local Similarity	23.3%;	Pred. No. 1.2;		
Matches	62;	Conservative 37;	Mismatches 116;	Indels 51; Gaps 8;
Qy	557	GGYKGKDYSEAEALIMTFSLN	PAGDPRLAQAQKWEAA--	FLEEMPAFORRMAG--MFOV 613
Db	120	GGLSGRDAAAGTAL--	---TVAQIPLOEDGKEAVARIRAYRVRVASPALDI 165	
Qy	614	TFTAERSLSEDEINRTTAED	-----LPFATSYVIVIELYVLSALGSYSSWSRVMVDSK 665	
Db	166	RVTGQQAIDDTTELAERADTKRSEFVALPLIALLLLVF	-----	204
Qy	666	ATLGLGGVAVVVLG--AVMAAMG--	FFSYLGIRSSLVLTLQVVPFLVLSVGADNIFIVLEY 721	
Db	205	GALVATGLPLVIGMLSTSVAMAGLYFLT	TWNEVSTPAQSVVTMLGLGAGIDYALLMWRNP 264	
Qy	722	QRLPRRPGEPREHVHIGALGRVAPSMLLCSLSEAI	CFPLGALTMPMPAVRTRFALTSLGLAVI 781	
Db	265	REELARRNAEAA--AARTVETAGRSVLPSGSTVAI	AMAGLLLPPLSPVRSIGIGVLAUV 323	
Qy	782	LDFLLQMSAFVALLSLDSKQREASRL	807	
Db	324	LTVLASITALPALITLLGDRVNA	PRL 349	

RESULT 72
D72482
probable antibiotic transport-associated protein APE2500 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: D72482
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahara, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawarabayashi, Y.
DNA Res. 6, 83-101, 1999
A>Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain K1
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72482
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1172 <RAW>
A:Cross-references: UNIPROT:Q9Y8Y4; UNIPARC:UIP000005E388; DDBJ:AP000064; NID:g5105945;
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE2500

```

Query Match      1.8%; Score 125.5; DB 2; Length 1172;
Best Local Similarity 18.2%; Pred. No.2,2;
Matches 140; Conservative 118; Mismatches 241; Indels 269; Gaps 32;

Qy 545 DYGAIPVPPFLAIGYGYKDYSEALIMTFSLNYPAGDPRLAQAKLWEBAFLIEMRAFQ 604
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 600 DMGGFLAYLGGGEG--AKAAEDVRRIIVNEALGDARVMLA--GEARLRKL----- 650

Qy 605 RRMAGMEQVWTFARRSLDEINRTTAEDLPFATSYIVIFYISLALGSYSMSRVMDVS 664
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 651 -----QEAVVEDLRRSDA-----ASSIIWAI----- 672

Qy 665 KATIGLGGAIVVLGANVMAAGFFSYLGRSLSVLTQLVVPFLVLSVGADNITPIFVLEVQR 724
   : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :

```

Db	673	-----LAAVIGTI-----LGV-----IIPFIGAG-----	693
Qy	725	PRRGPPEVHIGRALGRVAPSMLLCSLSAICFFLGA---LTPMPAVRTFALTSLGLAVI	781
Db	694	-----LTAAL---AIIYFLASNDLVDVTSQSRAIVFTTGLGLG	728
Qy	782	LDFFLQMSAFVALLSLDSKR-----QASRLDVCCCVKQOELPPPGQEGELLGFFQKAY	836
Db	729	IDY-----SAVYTKFRDINIRSGRGAEAGKAASNSVRP-----VLAGAFAAAA	773
Qy	837	A-----PFLHWTIGVVLLFLALPGVSLYSMCHISVCLQDELALPKDSYLLDY	886
Db	774	GFSSLMALAFPPFKVKSIGATVPAILLSVASISUTLTPAILALVGNWAGILWFFTCPY	829
Qy	887	FLFLNRYFEVGAQVYFVFTILGYNFSSEAGNNA-----ICSSAGCMNPSFTQKLOYAT	938
Db	830	-----KAYQPSRVP-----KSLVSAAGKAAFPVLLPLLVAAGLAG-----VYAA	869
Qy	939	EPPEQSY---LATPASWDDFDIMLTPSSCRLYISGPNKDKPCPSTVNSLMCLKNCMS	995
Db	870	TSEPGSDIILISIPRETGVYESMOHL-----VSEYDASRLFPQYIVA-----	911
Qy	996	ITWGSVRPSVEQHKYLPWFLNDRPNIKCPKGLAAVSTSVNLTSDGOVLASRFMAYHKP	1055
Db	912	SSSTLAPDV-----ADAADLDCVR-----RAETDGLLI-----YVV	943
Qy	1056	LKNSQDYTRALRAARELAANITADLRKVPCTGDPAFEPVPTITNVFEOYL-----TIL	1109
Db	944	LDTNPLNTEAVKVEELRG-----KYREVDPGSLVGGTAAVNLDRLDYLPFAYNRVI	996
Qy	1110	PEGLFMLSCLVPTF-AVSCLLLGL-DLRSGLLNLLSIVMILVDTVGFMAWMDISYNVAVS	1167
Db	997	PAAAVLVAIIAVFYGSIPAAAGVASILPAAAWSLSIVSILSGWTGIEAPW---PAPVL	1053
Qy	1168	LINLVSAVMSVFVSHITSPALSTKPTWLEPAKATI-----SMGSVAVFAG	1215
Db	1054	LAALLGVGM--DYISF-----YVNHAREAYLRGHSRYVEAASTGTGLVIG	1099
Qy	1216	VAMTNLPGLVLVLGAKAQLIQIPFFRLNLLITLLGLLHGLVFLPVILS	1263
Db	1100	LAMIMGAYAGLTLSRVEALQATGVSLSGVLGALLASLLIPFWA	1147
RESULT 73			
A:Probable RND efflux transporter PA3676 [imported] - Pseudomonas aeruginosa (strain)			
C:Species: Pseudomonas aeruginosa			
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004			
C/Accession: A83186			
R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, ; Lory, S.; Olson, M.V			
Nature 406, 959-964, 2000			
A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic			
A/Reference number: A82950; PMID:20437337; PMID:10984043			
A/Accession: A83186			
A>Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-1025 <STO>			
A/Cross-references: UNIPROT:Q9HKW4; UNIPARC:UPI000000C5AB4; GB:AE004787; GB:AE000			
A/Experimental source: strain PA01			
C/Genetics:			
A:Gene: PA3676			
Query Match 1.8%; Score 124; DB 2; Length 1025;			
Best Local Similarity 26.2%; Pred. No. 2.3;			
Matches 70; Conservative 40; Mismatches 109; Indels 48; Gaps 11;			
Qy	1017	NDRENIKCPKGLAAVSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTRALRAARELAANI	1076
Db	268	NDPPAPMRPMGEDAIGLAVAMKPGGDIL-----VLGKALE-----TEFARLQOOSLPAGL	317
Qy	1077	TADLRKVPGTDPAFEPVPTITNVFEOYLITLPEGLFMLSCLVPTFVAVSCLLGLDLR	1136

Db 317 IAGATFCLSPTRLPYFQTLGVPLAIGWIVVAAALTLGPALIAVTSRFGKLEPKRM--- 373
QY 811 CVKQOELPPGQEGELLGFFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVG 870
Db 374 -----ARVGRKVGAA--IVRW--PGPILGAVALLVGLTLPGYRTN 414
QY 871 LQELALPKDSYLLDYFLPLNRYFVGAPVYFVTLGYNFSSEAGHNAICSSAGCNSFP 930
Db 415 YNDRNYPADLPANEGYAAERHF-----SOARN----- 444
QY 931 TOKIOVATEPPRQSVLAIPASSWVDIFDMLTPSSCCRLYISGPNKDKPCPSTVNSLNC 990
Db 445 -----PE--VLWESDHDMENSADFLVINKIAK-----AIFAVEGI 478
QY 991 KNCMSITMGSVRPSVEQFHYLPWFNLDRPNIKCPKGGLAAYSTSVNLTSF--GOVLASRF 1049
Db 479 SRVQAITRDPGKP-IE--HTSIPFL-----ISMQGTSQLTEKYNQDLTARM 522
QY 1050 MAYHPLKNSQDYTEALRAELANIT-----AD----- 1079
Db 523 LEQVNDIQSNIDQMRHSLTQOMADVTHEMVIQTMGVVDVEELNHIADDFDPRPIR 582
QY 1080 -----LRKVGTDPAPFVFPVTLNV--FVEQVLTILPEGLFMLSCL 1120
Db 583 SYFVEKHCYDIPVCWSLRSVFTLDGIDVMTEDINNLPLMQRLDTLMPQLTAMPEMI 642
QY 1121 VPTFVAVCLLGL--DLRSGLNLLSIVMLVDTVGFMAWDISYNAVS----- 1167
Db 643 QTMKSMKAQMLSHSTQEGLODMAAQ--EDSAAWGEAFDASRNDSDSYLPEVPDNP 700
QY 1168 -----LINLVSAGVMSVEP-VSH-----ITRSPAISTKPTWLERAKEA---TISMGS 1210
Db 701 FORGLEQFLSPDGHAVRFITISHEGDPMSQAGIARIKTKT-----AAKEAIKGTPLSGS 754
QY 1211 AVFAG-----VAMTNLPGTLVLGLAKAQILQIFP----- 1239
Db 755 AYLGGTAMFKDLSDGNTYDMLIAGISALCLIFIMLITRSVAAAVIGTVVLSLGA 814
QY 1240 --FRLNLLI--TLGLH-LHGLV--FLFVILSYVGVDPNPALEQKRAEAAVAV 1287
Db 815 SPGLSVLIHQHILGIELHVLVLAAMAVIILLAVGADYN--LLLVARLKEEIHAGI 866
RESULT 76
G69154
conserved hypothetical protein MTH420 - Methanobacterium thermoautotrophicum (strain Del
C:Species: Methanobacterium thermoautotrophicum
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: G69154
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
Qi, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Ji, S.; Church,
G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
A:Reference number: A69000; MUID:98037514; PMID:9371463
A:Accession: G69154
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-676 <MTH>
A:Cross-references: UNIPROT:Q26520; UNIPARC:UPI000006649C; GB:AB000826; GB:AB000666; NID
A:Experimental source: strain Delta H
C:Genetics:
A:Gene: MTH420
Query Match 1.8%; Score 122.5; DB 2; Length 676;
Best Local Similarity 20.4%; Pred. No. 1.7;
Matches 104; Conservative 65; Mismatches 160; Indels 181; Gaps 24;
QY 472 YAPLNPNTSLYDCCINSILQYFQNNRTLLLTANOTLMGQTSQVMDKDHFLYCANAPIT 531
Db 27 YQPLMYDMDSYN-----LRLASNLNNEGELSEGWDRIYSYPPGVPL- 69
QY 532 FRDGTALALSCWADYCAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKL 591

Db 70 -----DY-PELLPHLTVLY-----LLFTWLL-----PGLTDTAP 98
QY 592 WEEAFLEEMRAPQRMAGMFQVTTAERSLE--DEINRTTAEOLDPIFATSY----- 640
Db 99 MLPALIAP-----LAGV--AVFAARSLDCDDSVATTAGLLATAAPFFWRTIPGFPD 149
QY 641 -----IVIFYISLALGYSYSSRSMVVDKATLGL-----GGVAVVLGVAAM 684
Db 150 TDMFNVLVPLAILLTLHLSLRDGW-----RMAVLSGALMGVFAAAWNGHQLISYIMALS 205
QY 685 GPFSYLGITRSSVILQVVPFLVLSVGAONIFIVLEYQELPRRCPGEPRHVIGRA----- 739
Db 206 ILHSLLYRRREPLI-----FLITS-----TLIVTMDPRAIPSPVGLMKIPVARADPPD 256
QY 740 -----LGRVAPSMLLCSLSEACIFLGLATPM--PAVTFEALTSLGLV 781
Db 257 PYANITELQRPQDDVLMALGPGLLLAGLGRFSIFRDWRDSNVLPVILWTLTGNASL 316
QY 782 -----LDFLMQSAFVALLSLDSKQRBASRLDVCCCKVQOELPPGQEGLLGF 831
Db 317 WGIREFSELLTAPLLITSALFLADLTASR-----VPP-----GF 350
QY 832 FQKAVAPFLHWTIRGVVLLFLALFGVLSYMSCHISVG---LDQ-----ELALPKDSYLL 884
Db 351 RRK-----LH-VAVAIMVLPFLIISTGQYSALHPRVDDGLLDAADYIRAKTPPTVVI 403
QY 885 D-----YPLFLNRYFEVGAPVVFVTTLGY 908
Db 404 SNWVGHFFAFMAR-----RPVNFQGRLAY 428
RESULT 77
Q69E11
ribonucleoside-diphosphate reductase (EC 1.17.4.1) large chain - human herpesvirus 4 (str
N:Alternate names: ribonucleotide reductase large chain
C:Species: human herpesvirus 4, Epstein-Barr virus
C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 09-Jul-2004
C:Accession: A03753; A22907; S32996
R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
Mol. Biol. Med. 1, 21-45, 1983
A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus
A:Reference number: A93065; MUID:85035713; PMID:6092825
A:Accession: A03753
A:Molecule type: DNA
A:Residues: 1-826 <BAM>
A:Cross-references: UNIPROT:P03190; UNIPARC:UPI000000C0CA; EMBL:V01555; NID:959074; PID:
R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.;
Nature 310, 207-211, 1984
A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
A:Reference number: A03794; MUID:84270667; PMID:6087149
A:Contents: annotation; protein coding region
R:Gibson, T.; Stockwell, P.; Ginsburg, M.; Barrell, B.
Nucleic Acids Res. 12, 5087-5093, 1984
A:Title: Homology between two EBV early genes and HSV ribonucleotide reductase and 38K
A:Reference number: A22907; MUID:84247360; PMID:6330697
A:Contents: annotation
C:Superfamily: herpesvirus ribonucleoside-diphosphate reductase large chain
C:Keywords: deoxyribonucleotide biosynthesis; early protein; oxidoreductase; redox-activ
F187-403,822-825/Disulfide bonds: redox-active #status predicted
F387,391/Active site: Asn, Glu #status predicted
F389/Active site: Cys (cysteine thiol radical intermediate) #status predicted
Query Match 1.8%; Score 122; DB 1; Length 826;
Best Local Similarity 20.4%; Pred. No. 2.4;
Matches 96; Conservative 46; Mismatches 150; Indels 178; Gaps 26;
QY 68 GDHLLILKIKCPRLVTGNTQACCSAKOLVSLASL-----SITKALLTRCP-----ACSD 118
Db 229 GKHSLLMRMI-NSHVEYHNYGC---KRPVSVAAMEPWHQIQFLETLKLPENHERCPG 284
QY 119 NFVNLHCHTCSPNOSLFINTVTRVAQLG-----AGQLPAVVAVYAFYOHSAEQSY 169

```
Db 285 IFTGLFV-----PELFKFLFRDTPMSDWYLFDPKDAGDL-----ERLYGEBFEREY 331
Qy 170 -----DSCSRVRPAAATLAVGTMCYVYVYGA-----LCNAQRWLNFGDGTGNGLAPL 216
Db 332 RLVTAGKFCGRVSTKSLMFISVN--CAVKAGSPFFILKEACNAHFWRDLQGEANNA----- 385
Qy 217 DITFHLLEPGQAVGGIPLNEGVARCNESQGDVATCSCDCAASCAPAIARPOALDSTF 276
Db 386 -----ANLCAEVLQPSKSVATCN-----LANICLPRCLVNAFLAVRAQADTQ- 429
Qy 277 YLQG-----MPG-----SLVLI--IILCSVFAVVTILVNGFRVAPARDKS 314
Db 430 --GDELLLALPRLSVTLFEGAGVGDGFSLARLDATQCATFVACSILOQSPTYDSRDM 487
Qy 315 KM-----VDPKGTSLSDKL-----SFSHTLLG----- 338
Db 488 SMGLGVQGLADVFDLQGWYTD-PSRSLNKEIFEHMYFTALCTSSLTGLHTRKIFPGFK 546
Qy 339 --QFQGWGTWASVPLTILVLSVIPVVALAAGLVFTLTDPVBLWSAPNSQARSEKAF 396
Db 547 QSKYAGGFHW-HDW-----AG---TDLST-PRINSRLSERIVROGLP 585
Qy 397 HDQH-----FGPFR-----TNQVILTAPNRSSVRY 422
Db 586 NSQFTALMPTSGCAQVTCGDAFYFPYANASTKVTNKEALRPNRSFWRH 635

RESULT 78
T52483
hypothetical protein [imported] - Thermus aquaticus
C:Species: Thermus aquaticus
C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004
C:Accession: T52483
R:Fee, J.A.; Chen, Y.; Barquera, B.; Teasos, P.; Gennis, R.
submitted to the EMBL Data Library, June 1999
A:Reference number: Z26089
A:Accession: T52483
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-385 <FES>
A:Cross-references: UNIPROT:Q9XK7; UNIPARC:UPI00000BC751; EMBL:L09121; PIDN:AAD38864.1
A:Experimental source: strain H88
C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein MTH76

Query Match 1.8%; Score 121.5; DB 2; Length 385;
Best Local Similarity 23.5%; Pred. No. 0.94;
Matches 77; Conservative 36; Mismatches 119; Indels 95; Gaps 13;

Qy 669 GLGGVAVVLGAVMAAMGFFSYLGIRSSLVILQVVPFLVLSVGADNIF-----IFV 718
Db 22 GPGTIGLVEGVAENTASLFKVVGGRSLDRIGRRRPFLLIGYGLPALFRPLLALAQSPHV 81
Qy 719 LEYQRLPR-----RPGEPREVHIGRALG-RVAPSMLLCSLSBAICFFLGAL 763
Db 82 LLYRFLDRTGKLRTPARDALTAESAPKEA-LGRAYGLHRSLDLTLCATLGPPLAFL--L 137
Qy 764 TMPAVRFTALTSGLAVIDLFLQMSAFVALLSLDSKQESRLDVCVCVPOELPPGQ 823
Db 138 LPLLGRGVFWLSALPALLAFLLLFAV-----RETFR-----PPSLPP--- 176
Qy 824 GGLLGLFPQKAYAPFLHLHWITRGVVLLFLALFGVSLVSMCHISVIGLDOELALPKDSYL 883
Db 177 ---LALHLSPGYRRF-----LLVSGLPALALSSNAFLRL-L-KELGLSQEVA 221
Qy 884 LDYFLF-----LNRYFEVGAPVYFVTTLGYNFSSEA-----GMN 917
Db 222 LAYTLVNLLYALLAYPLGGLADRVGLGRMVATFGGLYALVYLGFAWARTAFWALGFLEY 281
Qy 918 ALCSSAGCNFSTQIKIQVATEFPQS 944
Db 282 ALYSAA-----FEGANRAVLATLVPEEA 304
```

```
RESULT 79
H91123
probable oxidoreductase ECs3960 [imported] - Escherichia coli (strain O157:H7, substrain
C:Species: Escherichia coli
C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: H91123
R:Hayaashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
sasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A>Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: H91123
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <HAY>
A:Cross-references: UNIPROT:Q8XAL3; UNIPROT:Q8FDE8; UNIPARC:UPI00000D06C0; GB:BA000007;
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: ECs3960

Query Match 1.8%; Score 121.5; DB 2; Length 477;
Best Local Similarity 19.4%; Pred. No. 1.3;
Matches 90; Conservative 73; Mismatches 155; Indels 145; Gaps 23;

Qy 890 LNRFFEVG---APVVFVTTLCY-----NFSSSEAGMNA-ICSSAG----- 924
Db 29 INNIEGLASAPMFFLATIFYPFCLIIAEFVSLNKNSEAGVYAVYKSSILGGRWAFIT 88
Qy 925 -----CNFSFT----QKIQYAT-EPPEQSYLAIPASSWDD-----FIDWLTSSCCR 968
Db 89 AYTVMFVNLFFTSLLPRVIAYASYAFGLGYEYIMTPVATTIISMVLPFAFTWSTNGAKM 148
Qy 969 LYISGPNKDKCPSTVNSLCLNCKMCSITM-----GSVRP-----SVQFHKYLPW-FLN 1017
Db 149 L---GP-----ITSVTSTMLLLTSLYLLAGTALVGVQPADPTVDMAPNFWNAFLG 200
Qy 1018 DRPNIKCPKGLAAYSTSVNLTSDGQ-----VLASRFMAYHKPLKNSQDYTEALRAAR 1070
Db 201 VTTWIFMAGAESVAVYVNDVKGSKSFVKVILLAGIFIG-----VLYSVS 247
Qy 1071 ELAANITADLRKFGTDPAFEVF-----PYTITNVF----- 1101
Db 248 SVLINVFVSSKELKFTGGSVQVFGHMAAYFGLPEALMNRFFVGLVSFTAMFGSLLMWTATP 307
Qy 1102 YEQVLTILPEGLFMLSCLVETFAVSCLLGLDLRSLGLNLLSIV-MILVDTVGFMAWD 1160
Db 308 VKIFFSEIPEGIFGKK-----TVELNENGVPARAAWIOQLVIVLPLMIPLMGSNTVOD 360
Qy 1161 ISYNAVSLINLVSAVGMSEVF-----VSHITRSPFAISTKPTWLERAKEATISM 1208
Db 361 LMN---TIINMTAAASMLPPLFIMLAYLNLRKLDHLPRDFMGSRRTGI-----IVVSM 412
Qy 1209 GSAVPA-GVAMTNLPGLILVLGAKAQLIQIIPFRNLNLLITLG 1250
Db 413 LIAIFAVGVFASTFP-----TGANILTIIFVNVGGIVIFLG 448

RESULT 80
G85968
probable oxidoreductase ygjI [imported] - Escherichia coli (strain O157:H7, substrain EI
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: G85968
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glanier, J.D.; Rose, D.J.; Mayhe-
killer, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dinalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85968
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-477 <STO>
A:Cross-references: UNIPROT:Q8XAL3; UNIPROT:Q8FDE8; UNIPARC:UPI00000D06C0; GB:AE005174;
```

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: y9j1

Query Match 1.8%; Score 121.5; DB 2; Length 477;
Best Local Similarity 19.4%; Pred. No. 1.3; Indels 145; Gaps 23;
Matches 90; Conservative 73; Mismatches 155; Indels 145; Gaps 23;
QY 890 LNRYFEVG---APVYFVTTLYG-----NFSSEAGMNA-ICSSAG----- 924
DB 29 INNTELGLASAPMPPLATIFYPIPPCLIIAEFVSLNKNSEAGVYAVKSSLLGRWAFIT 88
QY 925 -----CNNESEF-----OKIQAT-EFPEQSVLAPASSWVD-----FIDMLTPSSCCR 968
DB 89 AVTYMFVNLFFFTSLRLPRVIAVSAYFAFLGYEYIMTPVATTIISVMVLFAPSTWSTNGAKM 148
QY 969 LYISGNKDKPCFSTVNSLNLCKNCSITM-----GSVRP-----SVEQFHKYLPW-PLN 1017
DB 149 L-----GP-----ITSVTSTMLLLTSYILLAGTALVGGVQPADPTIVDMINFNFWAFLG 200
QY 1018 DRPNIKCPKGLAAVSTSVNLTSDQ-----VLASREFMAYHKPLKNSQDYTEALRAAR 1070
DB 201 VTTWIFMAAGBAESVAVYVNDVKGSKSPVKVILLAGIFIG-----VLYSVS 247
QY 1071 ELAAMITADLRKVGTDPAFEVP-----PYITITNVF----- 1101
DB 248 SVLINVFVSSKELKFTGGSVQVFHGMAYFGLPEALMNRFGVLGVSFTAMFGSLLMTWATP 307
QY 1102 YEOYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSLNLISIV-MILVDTVGFMAWD 1160
DB 308 VKIFFSEIPEGFGKK-----TVELNENGVPARAATIQFLIVPLMIIPIMLGNTVQD 360
QY 1161 ISYNAVSLINLVSAGVMSVEF-----VSHITRSFAISTKPTWLERAKEATISM 1208
DB 361 LWN-----TIINMTAASMLPPLFIMLAYLNRAKLOHLPRDFMGSRRTGI-----IVVSM 412
QY 1209 GSAVFA-GVAMTNLPGILVLGLAKQLIQIOPFRNLNLTITLG 1250
DB 413 LIAIFAVGVASTFP-----TGANILTIIFYNVGGVIFILG 448
RESULT 81
A:acriflavin resistance protein d [imported] - Brucella melitensis (strain 16M)
C:Species: Brucella melitensis
C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AE3557
R:DelVecchio, V.G.; Kapratral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Eizer, P.H.; Hagius, S.; O'Callaghan, D.; Letes
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A:Reference number: AD3252; PMID:11756688
A:Accession: AE3557
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1024 <KUR>
A:Cross-references: UNIPROT:Q8YCZ5; UNIPARC:UPI00005846D; GB:AE008918; PIDN:AAL53624.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BMEI10382
A:Map position: II
Query Match 1.8%; Score 121.5; DB 2; Length 1024;
Best Local Similarity 28.1%; Pred. No. 3.5; Indels 43; Gaps 12;
Matches 65; Conservative 36; Mismatches 87; Indels 43; Gaps 12;
QY 1120 LVPTFAVSCLLGLDLRSLNLISIVMLVDTVGFMAWDISVNAVSLINLVSAGMSV 1179
DB 346 WVIVLAVSVSLG--LRAGFVVSLSIPLVLAITFLMSLMDISLQSVSLGALINLGLV 403
QY 1180 EBFVSHITRSFAISTKPTWLER-----KEATISMSGAVF---AGVAMTNLPGILVLGLAKA 1232
DB 404 D-----DAMIAVEMVMVARLEHGDIPINKAATVYVSHTAFPLMTGLT-VT- IAGFIPIGLNS 457

QY 1233 QLIIQIFF-----FRNLMLIT-----ILGLLHGLVFLPVILSYVGPVDPVDPALAEOKRAE- 1281
DB 458 QAGEYTTFTLVVIAVSLVSVWVAVLFAELLGVTFLP-----KKMKPH---EERKSRF 507
QY 1282 -BAVAAMVVASCPNHPSPSVSTADNIYVNHSPGSGKAGAGAIS-NFLPNNGR 1330
DB 508 PEAFSRVLLLSMRHKWTTIITIVLVFI-----SVFGMGFIERQFPQSDR 553
RESULT 82
A:probable efflux pump STY2719 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A:Note: this species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AE0816
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerston, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Parrar,
.; S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A:Authors: Party, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AE0816
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1037 <PAR>
A:Cross-references: UNIPARC:UPI0000059ACA; GB:AL513382; PIDN:CAD07711.1; PID:gl6503697;
C:Genetics:
A:Gene: STY2719
C:Superfamily: acriflavin resistance protein
Query Match 1.8%; Score 121.5; DB 2; Length 1037;
Best Local Similarity 18.9%; Pred. No. 3.5; Indels 333; Gaps 47;
Matches 171; Conservative 127; Mismatches 274; Indels 333; Gaps 47;
QY 533 KDGTAALALSCMA--DYGAPVPFPFLAIGYKGYDSEAEALIMTFSLNNYPA----- 581
DB 255 QDGSVKGLDVTAVELGAEKYDLS-----RFGNPGASGLGVKLAS 295
QY 582 GDPRLAQAQKLMEEAFLEEMRAF---QRRMAGMFQVTTAERSLEDEINRTTAEDLPIFA 637
DB 296 GANEMATAKLVLDRLNELAQVPHGLGYKIA--YETTSFVKASIIDVVKTL---LEATA 349
QY 638 TSYIVIFLYISLALGSSYSWSRVWDKATLGGVAVVLGVAWMAAGFPFVILGRSSVL 697
DB 350 LVFLVMYLFQ-----NFRATL-----IPTIAPVPMVLMGTFSVLYAFGYSI 390
QY 698 ILQVVPFVLVSVG--ADNIFIFVLEYQRLPRPG--EPREV---HIGRALGR-VAPSMILC 750
DB 391 NTLTFAMVLAIGLLVDDAI VVVENVERIMSEGLTPREATKSMGQIQGALVGTAMVLS 450
QY 751 SLSEACFPGLGALTMPAVRTFALTGSLAVILDFLOMSAFVALSLDLSKQSEASRLDVC 810
DB 451 AVFVPMAPFGG--TTGAIYRQFSITIVSAMVLSVLMILTPAL-----C 493
QY 811 CCVFPQELPPQGE-----GLLIGF-----FOKAYAPFL---LHWITRGVLL- 851
DB 494 ATL-----LKPLHKGEGHQGFQGFNRTFNRAERYEKGVAKLHLSRLWLIYVLLIG 549
QY 852 ---LFLAL-----FGVSLYSWCHISVGLDQELALPKDSYLLDVL-----F 889
DB 550 GMVFLFLRLPTSPFLQEDRGMTTSI-----QLPSGSTQQQLKVVKEVENVYFTHEKNNI 605
QY 890 LNRYFEVGA-----PVYFV-----TTLGYNF-----S 911
DB 606 MSVFTVSGSGGNGQNVARMFVRLKDWADRDPTTGSSFAIIERATKAPNQIKEARVPAS 665
QY 912 SEAGNNAICSSAGCN-----NFSFTQ-----KI 934
DB 666 SPPAISGLSSAGPDMELQDHAGHDALMAARDQLIBLAGKNSSLTRVHNGLDDSPQL 725

QY 935 QYATEFPQSYLAIP-----ASSWVDDFDLWLTSPSSCCRLYSIGPNKDKFCPS 982
DB 726 QIDIDORAQALGSIDINDTLOTAWGSSVYNDFMD---RGRVKVYVQAAKYRMLPD 782
QY 983 TVNSLNCILKNCMSITMGSRVPSVEQFHKLWFLNDRPNIKCPKGGGLAAYST-SVNLTS 1041
DB 783 DINL-----WYVRNK-----DCGMVFFSAFATSRMET 809
QY 1042 GOVLASRFWAYHKPLKNSQDYTEALRAARELAANITADLRKPGTDPAFEPFYITNVF 1101
DB 810 GSPRLRYNGY-----SAVEIVGEAAPGVS-----TGTANDVM----- 842
QY 1102 YEQYLITLPEGLFM-----LSLCLVPT-FAVSCILGLDLOIRSGLLNLSL---VM 1147
DB 843 -ESLVHQLPGGEGLEWTAMWSQERLSGAQAPALYALISLVVFLCL-AALYESWSPFESVM 900
QY 1148 ILV--DTVG-FMALW-----DISYNAVSLINLVSAGVMS-----VEFVSHIT-RSFA 1190
DB 901 LVVPLGVIGALLATWMRGLENDVYFQ-----VGLLTVIGLSAKNAITLIVEFANEMNQKHA 956
QY 1191 ISTKFTWLERAKEATISGSAVF-----AGVAMTNLPGILVL-GLAKAQLIQI 1237
DB 957 LLDATYASRQRLRPIILMTSLAFIPGVLPMTATSGAGSGSHAVGTGYMGGMISATILAI 1016
QY 1238 PFPERL 1242
DB 1017 FVPL 1021

RESULT 83
H83071
RND multidrug efflux transporter MexD PA4598 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83071
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: AB2950; UID:20437337; PMID:10984043
A:Accession: H83071
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1043 <STO>
A:Cross-references: UNIPROT:Q9HV19; UNIPARC:UPI00000C5D80; GB:AB0040873; GB:AB004091; NID
A:Experimental source: strain PA01
C:Genetics:
A:Gene: mexD; PA4598
C:Superfamily: acriflavin resistance protein

Query Match 1.8%; Score 121.5; DB 2; Length 1043;
Best Local Similarity 17.7%; Pred. No. 3.5;
Matches 236; Conservative 166; Mismatches 398; Indels 531; Gaps 58;

QY 134 SLFINVTR---VAOLGAGOLPAV-----VAYEAFYQHSFAEQSYDSCSRVVPAAATLAVG 186
DB 16 ALFISLAGLLVISKULPVAGVQPNVAPPQITITATYP-----GASAKVLVD 59
QY 187 TWCVVYGSALCNAQWLNFGQ-DTGNGLAPLDITTHLLFPG----- 226
DB 60 SVTSVLSESLNGAKGLLYFESTNNSNGTAEIVVTF---EPGTDPLAQVDVQNRLKKABA 116
QY 227 ---QAV-GSGIQ-----PLNEGVARCNESQGDVATCSQDCQCAASPAIAR 268
DB 117 RMPQAVLTQGLQVEQTSAGFLLIYALSYKEGAQRSDTTALGDYAAARNINNELRLPGVGK 176
QY 269 PQALDSTFYLGQWPGSLVLIILCSVFAVVITLLGVFVAPARDKSKMVDPKK---GTS 324
DB 177 LQFFSSE-----AAMRV-----WIDPQKLVEFGLS 201
QY 325 LSDKLSFSFTLLIGQFFQGGTGWASWPLTILVLSVIPVAL--AAGLVFTLTDDPVEL 382

202 IDD-----VSNAIRQNVQ-----VPAGAFGSAPGSSAQELFATLAVK 239
383 WSAFNSQARSEKAPHQHFQFPFRFTNOVILTAPNRSYRYDSLLIGPKNFSGILDLDLLL 442
240 GTLDDPQFGQVLRANEDSLVRLADVARLELKESYNISSRLNGPTVVGATQLSPGA 299
443 ELLE-----LQERLRHLQVWSPQARNISLQIDICVAPLNPNTSYL-----DCCNLSLQ- 492
300 NAIQTATLVKQRLAELSAPFPE-----DMQYSV--PYDTSRFDVVAIEKVHTLIEA 349
493 -----YFQNNRILL--LTANQTLMQOTSOVDKQHFLYCANAPLFFKQGTALAL 540
350 MVLVFLMFLQVRYTLPSIVVPVCLLGLTMMV-----YLLGFSVNMVMTFMGMVLAI 404
541 SCMADYGAPVFPFLAIGYKGYSEABALMTFSLNNYPAGDPRLAQAKLWEERAFLEEM 600
405 GILD-----DAIVVENVERIMA-BEGISPA-----EATVKAM 437
601 RAFORMMAGMFQVTTAERSLEDBINRTTABDLPI--PATSYIVIFLYISLAL----- 651
438 KOVSGAIVGITLV-----LSAVFLPLAFMAGSVGVYQQFSVSLAVSILFS 483
652 -----GSYSSWR-----VWVDSKATLGLGGVAV 675
484 GFALITFTPALCATLLKPIPEGHHEKRGFFGAFNRGFARVTERYSLLNSKLVARAGRFLM 543
676 VLGAVMAAMPFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGPREVH 735
544 VYAGLVAMIGYF-YLRUPEAFVPAEDLGVMVDV-----QLP--PGASR--- 584
736 IGRALRVAPSMLLCSLSEATCFFLGALTMPMPAVRTFA-----LTPGGLAVILDPLQ 787
585 -----VRTDATGELEERFLKSRVAVSFLIS 611
788 MSFVALLSLDSKQKQASRLVCCCKQKQELPPGQEGELLIGFPQKAYAPLHWHIRG 847
612 GFSP-----SQGDNAALAF-----PTFKDWSERG 636
848 VVLLFLALFGVLSVSMCHISVGLDQELALPKDSYLL----- 884
637 -----AQSSAAAEIATALNEHPALPDGTVMAVSPPPINGLNSGGFALRLMDR 684
885 -----DYPL-----FLNRYFE--VGAPVYFV-----TTLGYNFSSEAG 915
685 SGVGREALLQARDTLGEGIQTNPKFLYAMMEGLAEAPQLRLIIDREKARALGVSPETISG 744
916 -MNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFDLWLTSPSSCCRLYSIGP 974
745 TLSAAGSEVINDEFNAGRQQRVVIAEQG-----NRMTPESVLELYV--- 787
975 NKDKFCPSTVNSLNCILKNCMSITM--GSVRPSVEQFHKLWFLNDRPNIKCPKGGGLAAY 1032
788 -----FNAAGNLVPLSAFVSVKWEEGPV-----QLVRY-----NGYPSIRIVGDAAPGF 831
1033 STSVNLTSQGVLASRFWAYHKPLKNSQDYTEALRAARELAANITADLRKPGTDPAFEV 1092
832 ST-----GRAMA-----EMERLASQLPAGI----- 851
1093 FPYTITNVFYRQYLTI-LPEGLFMLSLCLVPTFAVSCILGLDLOIRSGLLNLSIVMIL-- 1149
852 -GYEWTGLSYQEKVAGQATSLFALAILV-----FLLLVALYESWSPISVLMVLIPI 903
1150 --VDTVGMALWDISYNAVSLINLVSAGVMS-----VEFVSHITRSPAISTKPTWLE- 1199
904 GAIGAVLAVWVGSMNDVYFVGLTIITIGLSAKNAITLIVEF-----AKELWEOG 952
1200 -RAKEATISMGSAVPAVAMTNLP--GILVL-----GLAKAQLIQIFFFFRLNLTLL 1249
953 HSLRDAATEAARLRFRPIIMTSMATLIGVILPALASGAGASQRAIGTGVIGMLSATEL 1012
1250 GLIHGLVFLPV 1260
1013 GVL-----FVPI 1019


```
QY 588 OAKLWEAFLEBEMRAFORRMAGMFOVTPTAERSLEDEINRTTAAEDLPFATSYIVIFLYI 647
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 676 PS-VNE-----SKVSPFLDSQOMLSLSVIOQCLX----- 705
QY 648 SLALGSYSSWRVMVDSKATGLGSAVAVLVGAVMAAMGFFSYLGIRSLVILQVVPFLVL 707
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 706 --AGKKQW-----RTASLTNWCAGL-----LAGLKVLPILVY 736
QY 708 -----SVGADNFIFFVLEYQRLPRPRGEPREHVHGRALGRVAPSMLLCSLEAI 756
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 737 KFFCLLTCNICIKQDSSMW-----KALHALRPOOL----- 767
QY 757 CFFLGALTPMFAVRTFAITSLGAVLDFLLQMSAFVALLSDSKRQEARLDVCCCKPKQ 816
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 768 -----TTEVSSGOAIFONILTEG-----DICASQRA 795
QY 817 ELPPPGOGEGLLGFFQKAYAPFLFHLHITRGVLLFLFALFGVLSYMSCHISVGLDQELA 876
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 796 AC---EGLGLLARLGNDFIARMI-WALHG-LALLTTEAAGLSFVSHVQAALGLALDIL 848
QY 877 LPKDSYLLDYFLFLNRYFEVGAPVYVVTILGYNFSSEAGMNAICSSAGCNPFSPTQKIY 936
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 849 LTESGWLDSQIGIRLINA-----IVAVLGPGLSPGSLFSCRKSV-IABISWQEIPT 902
QY 937 ATEPPEQSYLAIPASSWDDFDIMLTPSSCCRLYISGPNKDK-----FCPSTVNS 986
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 903 LLEYVNAFLFATILILHCYYMYNLYPVSSL-IFIVNLNIFRSVCFTQOQLILFAPQAVSV 961
QY 987 LNLCLNCSITMGSRVPSV-----EQFKYLPWFLLND----- 1018
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 962 HHVKNLL-MTLASQPIIRLSVSTRHLVKEKDPVSVIDEQIEDNLQMLDEETDSBIG 1020
QY 1019 -----RPN-----IKCPKGLLAYS-----TSVNLTSQO 1043
Db 1021 NLIRSTLIRLLYATCPSRPSRMMLICRNMAAASAGSAETSIAENDPAYTRENLGDDDE 1080
QY 1044 VLASRFMAVHKPKNSQDVTEALR-AARELAANITADLRKVGTDPD-PEV----- 1092
Db 1081 DMVS--SSGKSIIRANPKDKTLRTRVFAAECLSLLEPAVGNDAHFIDILLARNLASN 1138
QY 1093 -----PPYITNVFYEQYLFILPEGLFMLSCLVLPFTFVASCLLGLD 1135
Db 1139 RQSGDWLVQLQELISLAYQISTI---QFENMRPIGVGLLSTILEKAQLLSAVRTALDA 1195
QY 1136 RSG 1138
Db 1196 NSG 1198

RESULT 86
AF3386
protein translocase chain secD / protein translocase chain secE [imported] - Brucella me
C:Species: Brucella melitensis
C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C:Accession: AF3386
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Loeb, T.; Ivanova,
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hegius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A:Reference number: AD3252, PMID:11756688
A:Accession: AF3386
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-775 <KUR>
A:Cross-references: UNIPROT:Q8YGT1; UNIPARC:UPI0000057FID; GB:AE008917; PIDN:AAL52257.1;
A:Experimental source: strain 16M
C:Genetics:
A:Gene: BME11076
A:Map position: 1

Query Match 1.7%; Score 119.5; DB 2; Length 775;
Best Local Similarity 21.0%; Pred. No. 3.3;
Matches 120; Conservative 88; Mismatches 223; Indels 141; Gaps 27;
```

```
QY 334 HTLQGPFQCGMT---WVASWPLTILVLSVTPVVALAAGLVFETLTTPDVELWSPAPNSQA 390
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 16 HSMU--YFSRWKSAIWLIA-----VLVSLII-----ASPNFLS 46
QY 391 RSEKAFHQHFGPPFRTNQVILTAPNRSSRYSDLSLLGPKNFSGILDLDLLELELQER 450
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 47 RETLA-----NLPDFLPKKQVSLGLDLSGGR--LILQVQN-AGKTDLTTANI--MRQR 96
QY 451 LRHLQWNSPEAQ---RNISLQDICYAPLNPNTSILYDCCINSLLQYQNNRTLLLTANQ 507
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 97 LEELGYGNPVVEGGRN-----QIRVEVPGLYDA-----QLLKDILITIRNL 138
QY 508 TLMGQTSQVDWKO-----HFLYCANAP-----LTFKDGTTALALSCMADY 546
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 139 SFRAMDDTMSDDDAIRGTPLADSEIVYFDDPPVGYLLKKTPILTGHDITDAKASISADD 198
QY 547 GAPVFPFL-----AIGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLM-EEAF-- 596
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 199 GQPVITLTDNGRRRLADLTAQGNENSAFIVVDNQVVSAPTSGPLDTSBLQIEGAFDL 258
QY 597 --LEEMRAFQRMAGMQVTFTAERSLEDEINRTTAAEDLPFATSYIVIFLYISALGSY 654
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 259 QAANNMAVLRSGALPOAVTVLEERTIASAL---GED---YASAAVLAALLAALVVGULF 311
QY 655 SSWSRVMVDSKATGLGSAVAVLVGAVMAAMGFFSYLGIRSLVILQVVPFLVLSVG-ADN 713
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 312 MVLSYGLIGVIALVALVNNIILTAVLSLIG-----ASISLASTAGLVLTIGLAVD 362
QY 714 IFIFVLEYQRLPRRP---EPREVHIGRALGRVAPSMLLCSLSEAI-CFFIGALTMPMAV 769
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 363 AHILIYERVRDRKRGYSVQWMSGFYRALSTIVDANLTLIAALVFLGSGT-----V 418
QY 770 RTPALTSLGAVLDFLLQMSAFVALLSDSKRQEARLDVCCCKPKQLPFGQEGELL 829
   :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 419 HGFALTVAIGTGTTLTTLT-FTRLIAQWVR-----TAKPKEVP-----KRLIK 462
QY 830 GFFQKAVAPFL-LHWITRGVLL---LFLALF 857
Db 463 LVPVTTHIPFMRLOFVLIGISVLACAVVALF 494

RESULT 87
S76433
cation efflux system protein czcA-2 - Synecchocystis sp. (strain PCC 6803)
N:Alternate names: protein sl10142
C:Species: Synecchocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: S76433
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecchocystis
s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76433
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1075 <KAN>
A:Cross-references: UNIPROT:P74461; UNIPARC:UPI00000D3597; EMBL:D90915; GB:AB001339; NID
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Genetics:
A:Gene: czcA-2
C:Superfamily: acriflavin resistance protein

Query Match 1.7%; Score 119.5; DB 2; Length 1075;
Best Local Similarity 20.2%; Pred. No. 5.1;
Matches 239; Conservative 140; Mismatches 417; Indels 385; Gaps 58;

QY 353 LTILVLSVTPVVALAAGL-----VFTELTPDVE-LWSAPN--- 387
Db 26 LAIIVLGVFAVFSLPVDLPSITYPRIGVRLDAPGVSPVAVDTEITRPLEAALSATEGV 85
```


QY 1172 ----VSAGVMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227
Db 961 IEATLDVAVMRRLPILMTSLAIFLGMP-----LVISTG-----AGSGAQNAGVTGVM 1008
QY 1228 -GLAKAQIQUIFF 1239
Db 1009 GGMVATVLAIF 1021
RESULT 95
C90693
acridine efflux pump [imported] - Escherichia coli (strain O157:H7, substrain RIMD 05099)
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: C90693
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: C90693
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <HAY>
A:Cross-references: UNIPROT:Q8XD55; UNIPARC:UPI00000D03CA; GB:BA000007; PIDN:BA033938.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: ECs0515
C:Superfamily: acriflavin resistance protein
Query Match 1.7%; Score 117.5; DB 2; Length 1049;
Best Local Similarity 17.7%; Pred. No. 6.8;
Matches 140; Conservative 117; Mismatches 249; Indels 287; Gaps 33;
QY 624 EINRTTADLPFAISYIVFLYISALGYSYSSWRVMDSKATL--GLGGVAVVLG--A 679
Db 339 EVVKTLVE-----AIIIVFLVYLFQNF-----RATLIPTIIVPVLGTF 381
QY 680 VMAAMGFFSYLGRISLVILQVVPFLVSVG--ADNIFVFLVLEYQRLPRRPGEPREVHIG 737
Db 382 VLAAGF-----SINTLTFMG-VLAIGLLVDDAIVVENVERVMAEGLPPKEATR 432
QY 738 RAIGR-----VAPSMLLCSLSBAICFFLGALTPMPAVRTALTSLGLAVILDFLLQMSAPV 792
Db 433 KSMGQIQGALVGIAMVLSAVFVPMAPFGSGTGAI--YRQFSITIVSAMLSVLVALILTP 490
QY 793 ALLSLDSKQERASRLDVCCVQQLPPGQEGILLGFFQKAYAPFLHVI-----844
Db 491 ALCA-----TMLKPIAKGDHGEKGGKGFNGFNMFPEKSTHHTYDTSVGGILR 536
QY 845 TRGVLLFLALFGVSLY-----SMCHISVGLDOE-----874
Db 537 STGRVILVLYIIIVGMAYLVFLVPLSPFLPDEQGVFMVWQLPAGATQRTOKVNEVTH 596
QY 875 --LALPKDS--YLLDYFLFNR-----892
Db 597 YLTKEKNVSEVFAVNGFGFAGRGQNTGIAFVSLKMDADRGEENKVEAITMRATRAFS 656
QY 893 -----YFEVGAPYVFTTLYGNFS--SEAGM-----NAICSSAGCNFSFT--931
Db 657 QIKDAMVAFNLPFAIVELGTATGDFDELIDQAGLGEKLTQARNQLLAAEAKHPDMLTSV 716
QY 932 -----QKIQVATEPEQSYLAIPAS-----SWVDDFDLWLTSSCRLVI 971
Db 717 RNLGLEDTPQKIDIDQEAQALGVSINDINTLGAAGGSYVNDIFID--RGRVKVTV 773
QY 972 SGNPKDKFCSTVNSLNLCKNCSMTGWSVRPSEQFHKYLFWFLNDRPNIKCPKGLAA 1031
Db 774 MSEAKYRMLPDDIGD-----WYVRAADQGVMP---FSA 803
QY 1032 YST-----SVNLTSDGQVLASRFMAHYKPLKNSQDYETALRAARELANITADLRKVP 1085
Db 804 FSSSRWEYGSPLRYNGLPMSIILQQAAPGKST---GEAMELMEQLASKLPTGV-----855

QY 1086 TDAPEFVPPYITITNVFYEQYL-----TILPEGLFMLSCL-----VPT 1123
Db 856 -----GYDWTGMSYQERLSGNQAPSLYAIASLIWVFLCLAALYESWSIPFSVMLVWEL 907
QY 1124 FAVSCLILGLDLRSGLNLLSIVMLIVDTVGPMAWLDISYNAVSLINL-----1171
Db 908 GVIALLAA--TFRGLTNDVYFQVGLLTTIGLSA-----KNAILIVEFAKDLMDKEGKGL 960
QY 1172 ----VSAGVMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227
Db 961 IEATLDVAVMRRLPILMTSLAIFLGMP-----LVISTG-----AGSGAQNAGVTGVM 1008
QY 1228 -GLAKAQIQUIFF 1239
Db 1009 GGMVATVLAIF 1021
RESULT 96
B36938
acriflavin resistance protein acrB - Escherichia coli (strain K-12)
N:Alternate names: probable transmembrane protein acrE
C:Species: Escherichia coli
C:Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C:Accession: B36938; E64776
R:Ma, D.; Cook, D.N.; Alberti, M.; Pon, N.G.; Nikaido, H.; Hearst, J.E.
J. Bacteriol. 175, 6299-6313, 1993
A:Title: Molecular cloning and characterization of acrA and acrE genes of Escherichia coli
A:Reference number: A36938; MUID:94012493; PMID:8407802
A:Accession: B36938
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1049 <MAA>
A:Cross-references: UNIPROT:P31224; UNIPARC:UPI0000125397; GB:U00734; NID:9392830; PIDN:
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E64776
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1049 <BLAT>
A:Cross-references: UNIPARC:UPI0000125397; GB:AE000152; GB:U00096; NID:gi786660; PIDN:A.
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: acrB; acrE
C:Function:
A:Description: multidrug efflux pump
A:Note: active form contains proteins acrA and acrB
C:Superfamily: acriflavin resistance protein
C:Keywords: transmembrane protein; transport protein
F:12-28/Domain: transmembrane #status predicted <TM1>
F:343-359/Domain: transmembrane #status predicted <TM2>
F:370-386/Domain: transmembrane #status predicted <TM3>
F:397-413/Domain: transmembrane #status predicted <TM4>
F:442-458/Domain: transmembrane #status predicted <TM5>
F:472-488/Domain: transmembrane #status predicted <TM6>
F:541-557/Domain: transmembrane #status predicted <TM7>
F:875-891/Domain: transmembrane #status predicted <TM8>
F:899-915/Domain: transmembrane #status predicted <TM9>
F:975-991/Domain: transmembrane #status predicted <TM10>
F:1012-1028/Domain: transmembrane #status predicted <TM11>
Query Match 1.7%; Score 117.5; DB 2; Length 1049;
Best Local Similarity 17.7%; Pred. No. 6.8;
Matches 140; Conservative 117; Mismatches 249; Indels 287; Gaps 33;
QY 624 EINRTTADLPFAISYIVFLYISALGYSYSSWRVMDSKATL--GLGGVAVVLG--A 679
Db 339 EVVKTLVE-----AIIIVFLVYLFQNF-----RATLIPTIIVPVLGTF 381
QY 680 VMAAMGFFSYLGRISLVILQVVPFLVSVG--ADNIFVFLVLEYQRLPRRPGEPREVHIG 737

```
Db 382 VLAAGF-----SINTLTFG-WLAIGLLVDDAIVVENVERVMAEGLPKEATR 432
Qy 738 RALGR-----VAPSMLLCSLSBAICFFLGALTPMPAVRTALTSLGLAVILDFLLQMSAFV 792
Db 433 KSMGQIQGALVGIAMVLSAVFVPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALILTP 490
Qy 793 ALLSLDSKQEASRLDVCCVCPQELPPPGQEGLLGFFQKAYAPFLHMI-----844
Db 491 ALCA-----TMLAPIAKGDHGEGKGFPGFNFNMFSEKSTHYTDSVGGILR 536
Qy 845 TRGVLLLFALFGVSLY-----SMCHISVGLDOE-----874
Db 537 STGRVLVLIIVGWGAYLVFLPSSFLPDEBQGVMTWVQLPAGAQERTQKVLNEVTH 596
Qy 875 -LALPKDS-----YLLDPLFLNR-----892
Db 597 YYLTKKNVSEVFAVFGFAGRGQNTGIAFVSLKMDADRFGEEKVCAITMRATRAFS 656
Qy 893 -----YFEVGAPYFVTTLGYNFS--SEAGM-----NAICSSAGCWNFSFT--931
Db 657 QIKDAMVFAFNLPAIVELGTATGPFELIDQAGLHEKLTQARNQLLBAAKHPDMLTSV 716
Qy 932 -----QKIQTATEPEOSYLAIPAS-----SWDDFDIDLTPSSCCRLYI 971
Db 717 RPNGLDTPQKIDIDQEKAAQALGVSINDINTLGAAGGVSYNDFID---RGRVKYVY 773
Qy 972 SGPNDKCPSTVNSLNCCKMCKSITMGSVRPSVEQFHKYLPWFNLDRPNKCPKGGGLAA 1031
Db 774 MSEAKYRMLPDIDG-----WYRAADGQWVP---FSA 803
Qy 1032 YST-----SVNLTSDGQVLAGRMAVYHKPLKNSQDYTEALRAARELAANITADLRKVP 1085
Db 804 FSSRWEGSPLERYNGLPSEILGQAPGKST---GEAMELQGLASKLPTGV-----855
Qy 1086 TDAPEVFPYTTINVFYEQYL-----TLPEGLFMLSCL-----VPT 1123
Db 856 -----GYDWTGMSYQERLSGNQAPSLAIYSLIVFLCLAALESNSIPPSVMLVPL 907
Qy 1124 FAVSCLLGLLGLSLGALLSIVMLTVGVFMALWDISYNAVSLINI-----1171
Db 908 GVIGALLAA--TFRGLTNDVFPQVGLLTIGLSA-----KNAILIVEFAKOLMDKEGKL 960
Qy 1172 ----VSAGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAFAGVATNLPGLVL 1227
Db 961 IEATLDVVRMLRPILMTSLAIFLGVP-----LVISTG-----AGSGAQNAGVTGM 1008
Qy 1228 -GLAKAQLIQIFF 1239
Db 1009 GGMVATVLAIPF 1021

RESULT 97
S23756
C:RTR protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 05-Mar-1994 #sequence_revision 01-Sep-1995 #text_change 09-Jul-2004
C:Accession: S23756
R:Tucker, S.J.; Tannahill, D.; Higgins, C.F.
Hum. Mol. Genet. 1, 77-82, 1992
A:Title: Identification and developmental expression of the Xenopus laevis cystic fibrosis
A:Reference number: S23756; MUID:93244789; PMID:1284470
A:Accession: S23756
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1485 <TUC>
A:Cross-references: UNIPROT:P26363; UNIPARC:UPI0000127558; EMBL:X65256; NID:g64622; PIDN
C:Superfamily: cystic fibrosis transmembrane conductance regulator; ATP-binding cassette
C:Keywords: ATP; nucleotide binding; P-loop; transmembrane protein
F:442-623/Domain: ATP-binding cassette homology <ABC1>
F:459-466/Region: nucleotide-binding motif A (P-loop)
F:1230-1422/Domain: ATP-binding cassette homology <ABC2>
F:1247-1254/Region: nucleotide-binding motif A (P-loop)
```

```
Query Match 1.7%; Score 117.5; DB 2; Length 1485;
Best Local Similarity 17.3%; Pred. NO. 11;
Matches 183; Conservative 156; Mismatches 375; Indels 343; Gaps 46;

Qy 265 AIARPQALDSFYLGQMPGSLVLI-----ILCSVFAVVTILLVFRVAPADKSKM 316
Db 558 SLARAVYKADLYLLDSFYSYLDLFTKEIFEESCVCKMANKTRILVTSKVEQLKRAKV 617
Qy 317 VDPKKG-----TSLSD-KLSFSTHTLLQFFQGMGTWVASWPLTILVLSVIPVALA 367
Db 618 LILHEGSCYFGTFSELEDQRPFSHLIGPDHEN-----AER 655
Qy 368 AGLVPE-----LTTDPVELWSAPNSQARSEKAFHDHFGFPFTNQVILTPNRSSVR 421
Db 656 RNSIITETLRCSIDSDP-----SAVRNEVKK-KSF--KQVADFTKRSKSIINPKSRK 708
Qy 422 YDLSLLGPKNFGSILDLLELLELQBRRLHLOWSPDAQRNISLQDICVAPLNPNTS 481
Db 709 FSLMQKSQPMQSGIEEDMPAEQGE-----RKLISVPESEQGEASLPRSNFLTGT---760
Qy 482 LYDCINSLLOYFQNNRTLLLTANTQTLMGOTSDQWMDHFLYCANAPLTPKOGTALAS 541
Db 761 -----FQGRRRQSVL---NLMTRTS-----ISQGS---N 783
Qy 542 CMADYGAPVFPFLAIGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWE-----593
Db 784 AFATRNASVRK-MSVNSYSNS-----SFDLDIY---NRRLSQDSILEVSEINBE 829
Qy 594 ---BAFEEEMAFQRRMAGMFQVTF--TAERSELEBEINRTTAEDLPFATSPVIVPLYSI 649
Db 830 DLKCEFLDDTDSQPTTWTNYTLAFLTAHKNF-----IFILVCLVTFPVEV 876
Qy 650 ALGYSYSSW-----SRMVDKATLGLGVAVVLGAVMAAMG-----FFSYLGRSSILVILQV 701
Db 877 AASSAWLMIKRNAPAINMTSNENVSEVSDTSLVIVTHTSFYVYFVYVGVADSLALGI 936
Qy 702 ---VPFLVLSVGADNIFIVLEYQRLPRRPEPREVHIGRALGRVAPS-----MLL 749
Db 937 FRGLPLVHSLISVSKVLHKMLHALHAPMSTFTMRAGRILNRPFKDTAILDILPLSI 996
Qy 750 CSLSEACIFFLGALTMPAVRTALTSLGLAVILDFLLQMSAFVALLSLDSKQREASRLDV 809
Db 997 PDLAQLVLIIGAITVSLLEPYIFLATVPVIVAFILLRSYFL-----HTSQKLQLE- 1049
Qy 810 CCCVKPQELPPPGQEGLLGFFQKAYAPFLHMIT--RGVLLLFALFGVSLY--SMC 865
Db 1050 -----SKARSPFPAHLITSLKG---LWTLRAFGRPYFETLP 1083
Qy 866 HISVGLQELALPKDSYLLDYFLFN--RYPEVGAPVVFVTTLGVNFSSEAGMAICSSA 923
Db 1084 HKALNL-----HTANWFLYLSTLRWFQMTIEMIFVI-----1114
Qy 924 GCNNFSPQKIQTATEPEQSYLAIPASSWVDDFDIDLTPSSCCRLYISGNKDK-----978
Db 1115 -----FFIAV-----SFISIAT-----SGAGEKVGIVL 1138
Qy 979 -FCPSTVNSLNCXCMISITMGSVRPSVEQFHKYLPWFLNDRPNKCPKGGLAAYSTSVN 1037
Db 1139 TLAWNIMNTLQWAVNA-SIDVDSLMSRSVSRIFRI-----DLP-----VEELIN 1181
Qy 1038 LTSQGVQLASRFMAYHKPLKNSQDYTEALR-----AARELAANITADLRKVPGTDA 1089
Db 1182 ENKKEEQLESEVLIY-----ENDYVKTKQWVPSGGQWTKVLSANY-----IDGGNTV 1229
Qy 1090 FEVFPYTTINVFYBQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSLNLSIVMIL 1149
Db 1230 LE-----NIFSLSLPQQRVG--LLG-RTSGKSTLLSAFURL 1263
Qy 1150 VDTVGFMAJWDISYNVSLNLSVAGMSVEFVSHITRSPAISTKPT--TWLERAKEATIS 1207
Db 1264 LSTQGDIIQDGVSWQTIPLQKWRKAFGVIQKVPFIFSGIRKNLDIPYKGV---SDEELLK 1320
```

QY 1208 MGSAPVAGVAMTNLPGIL-----VLGLAKAQLI 1235
DB 1321 VTEVGLKLIIDQFFQQLDVLDDGCVLGHGKQLV 1357

RESULT 98

B98137
hypothetical 46.1K protein in ploc 3'region [imported] - Agrobacterium tumefaciens (stra
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: B98137
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollan, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: B98137
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-631 <KUR>
A:CROSS-references: UNIPROT:Q8U610; UNIPARC:UPI00000D2701; GB:AE007870; PIDN:AAK88620.1;
C:Genetics:
A:Gene: AGR_L113
A:Map position: linear chromosome

Query Match 1.7%; Score 117; DB 2; Length 631;
Best Local Similarity 20.4%; Pred. No. 3.7;
Matches 92; Conservative 65; Mismatches 175; Indels 120; Gaps 17;

QY 500 LLLLTANQTLMOQSDVQDKHPLCYANAPLFKDGTA-LALSCMADYGAPV----- 550

DB 46 LLAMMIGLLLVGTSR-----YALHLPYIWDIETSLCFLWFMALGAAIAIDRSEHL 97

QY 551 -----PP-----PLAIGYKGYSEAEALIMTSLN----- 577

DB 98 RLTLFLNMEPQVILGYNSLALVATFLAALIKPAMEYAEIWEVWTSALNIPMSFRAA 157

QY 578 NYPAGD---PRLAQAKLMEAPLEEMRAFQRMAGMFQVTFTAERSLEDEINRTAEDLP 634

DB 158 ALPVGACLLMLLVNLPRNLRDIIAFTVVAAGLLYLASPMLESIGNLNLAIPLG 217

QY 635 IPATSYIVPLYSIALG-----SYSSW-SRWVWDSKATIGLGG-----VAVVIGA 679

DB 218 LFVAVPLVGVPIAFGLGTGLTAYLTFTTWTPTIVMIGRMDGMSGIILLSVFVLLGC 277

QY 680 VMAAMG-----PFSYIGIRSSVLIVQVPLVLSVG-----ADNIPFVLEYQRL 724

DB 278 VLDATGMGKAIVELLSMPFGHIRAGMSYVLLGSL-FLVSGISGSKVSDMATVAPALFPEM 336

QY 725 PRPGPREVHIGRALG-----RVAPSMLLCSLSEALICFFLGALTTPMPAVRTFALTSLGA 779

DB 337 KRGHKPKEMIALLATGAAMADTVPPSIVLVLVLSAAGVSIAGL-----FTSGPM 386

QY 780 VILDFLLWMSAFVALLSDSKQERLDCVCCVQPELPPPGQEGLLGLGFQKAYAPF 839

DB 387 IAMVLLLVLAFLARWAKDENMEGAKR-----TPVQVQVGAALIA-APALVLPF 434

QY 840 LLHMITRG-----VLLLLPLALPGVSLY 862

DB 435 LIRSLVGGGVATATEVSTIAVLYAMIVGAVLY 466

RESULT 99

T38495
hypothetical protein SPAC29B12.07 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T38495
R:Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21797
A:Accession: T38495
A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
A:Residues: 1-1969 <GEN>
A:CROSS-references: UNIPROT:O14029; UNIPARC:UPI000017B1B4; EMBL:Z99164; PIDN:CAB16252.1
A:Experimental source: strain 972h-; cosmid c29B12
C:Genetics:
A:Gene: SPDB:SPAC29B12.07
A:Map position: 1
A:Introns: 664/2

Query Match 1.7%; Score 117; DB 2; Length 1969;
Best Local Similarity 18.1%; Pred. No. 17;
Matches 204; Conservative 117; Mismatches 390; Indels 418; Gaps 46;

QY 92 SAKQIVSIEASISITKALLTRCPACSDNFVNLHCHTNCSPNOSLFINVTRVLAQLGAGQLP 151

DB 806 SAMSPGNLQRTANLYKPMTT---PNAYNIKSNQRETKYPYQPAINYSEVTPQSGSSLP 862

QY 152 AVVAYEAFQHSFPAQSYDSCSRVRVPAATLAVGTMCGVGSALCNAQRLN----- 204

DB 863 T-----SGERANIIRSPGFTPLAAQKDATIYTPSHAQATLYGNMNDNRD 907

QY 205 -----PQDGTGNGLAPLIDITFHLLEPGQAVGSG-----IQPLNEGVARC---NESQG 248

DB 908 NEGIIHDILOSDMEPVLPNNSAYHANAPVSSHSEGLNNLPLPISPLPQLHKTGTSHQHG 967

QY 249 DDVATSCODCAASCAPAIARPOALDSTFYLGMGPGSLVLIILCSVFAVVTILLVGRVA 308

DB 968 PDAETATTAQVAPSIPTPNPNVSDITMTGVALPSATL----- 1006

QY 309 PARDKSKVMDPKKGTSL-----DKLSFSTHTLLGQFPQGWGTWASWPLTILVLSVI 361

DB 1007 ---DSDKSLSLHKRSALSRNNSPRPDFLPLPNQPLH-----SNLHSPVS 1048

QY 362 PVVALAAGLVFTLTTPDVELWSAPNSOAR-----SEKAFHQHGFPPFTNQVILTAP 415

DB 1049 PV-----DSEDSRLKPLSTORPAP---SFGPC---GTIVMAP 1081

QY 416 NRSSRYDLSLLGPKNFGIILDLLELLELLEQLRHLQVMSPEAQRNLSLQDICYAPL 475

DB 1082 STPSGLYTTSGKGTFFIAGPIKIGLDV--LTDEYRHLKBFKG----- 1123

QY 476 NPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMOQSDVQDKHPLCYANAPLFKDG 534

DB 1124 -----PYLASN-----GKVDKHKGAIEWLSKYIDRLAQSLSEYDDK 1160

QY 535 -----GTALALSCMADYGAPVPPFLAIGYKGYSEAEALIMTSLNPNYPAGDPRLAQA 569

DB 1161 NITLKKKULLLOCL-----KMLEVS-----DRKL----- 1185

QY 590 KLWBEAPLEEMRAFQRMAGMFQV-----TPTAERSL-EDEINRTTAEDLPFATSYIVI 643

DB 1186 -----IVEKLKRP1---LLPSPEIPEPCNTATSVQELINPEINQ---DDSPIVASRYCTT 1233

QY 644 -FLYTSLSLALGYSYSSRWVWDSKATIGLGVAVLVGAVVAMGPFPSYLGIRSSVLQV 702

DB 1234 SPLH-----RFYEYL----- 1243

QY 703 PFLVLSVGADNIFIFVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALICFFLGA 762

DB 1244 ---LSGNKDEALTVALQOKOMP-----YAIIVAHSIDA 1273

QY 763 LTPMEPAVRTFA-----LTSGLAVILDFLIQMSAFVALLSDSKQERLDCVCCVCK 814

DB 1274 KTFQGVVTFCKSEVKESMLRSGVGNLSLSLQMSDAHASSMSRFSSTSLN----- 1327

QY 815 POELPPPGQEGLLGLGFQKAYAPFLHWTIR-----GVVLLFLALFGVLSVSMCHI 867

DB 1328 ---LADQASNALVAMKELLYNIIANHYSDQKRALVGLTLLQENRVAAHLYVLSL 1384

QY 868 S-----VGLDOELALPKDSVILDDYF-----LFNRYPEVGVAPVFTVL 906

DB 1385 SPDVCNSKNSLFLVGLSKHNLVPSHDDLFDVDTQTEVLELVFNVYS-KTFVFTFLV 1443

Qy 907 GYNE-----SSEAG-----MNAICSSAGCNFSFTQKIYATEPPEQSYL 946
Db 1444 PYRLYEAEVLAAGEVSAARKYCELYGNLRVAKSNVDPGFLVRDLT-----QQIL 1499
Qy 947 AIPA-----SSWVD-----DFIDWLTSSCCRLYISGPNKDKFCPSTV-----NSL 987
Db 1500 ENSAGSEDISSWLGRTVSRPLDVLSSLGSKFSKFVAGDPNFDVMPATVGPFGKV 1559
Qy 988 NCLKNCMSITGWSVRPSVEQPHKLPWFNLDRPNIKCPKGGGLAAYSTSVNLTSDGOVLAS 1047
Db 1560 ASQKN-LTVQNTNNAAMESPY-----SDRPTSSGP-----SYQNRTPLT--GQ--ES 1602
Qy 1048 RFMAVHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYT 1096
Db 1603 MMGVSPYRRS-----TEIAENMMDGNAYPYT-PASQENPYT 1640

RESULT 100
AE0380
multidrug efflux protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AE0380
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AE0380
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1050 <KUR>
A:Cross-references: UNIPROT:Q8ZC87; UNIPARC:UPI00000DCB3F; GB:AL590842; PIDN:CAC92368.1;
C:Genetics:
A:Gene: acrB
C:Superfamily: acriflavin resistance protein

Query Match 1.7%; Score 116; DB 2; Length 1050;
Best Local Similarity 17.7%; Pred. No. 8.7; Mismatches 243; Indels 310; Gaps 35;
Matches 144; Conservative 115;

Qy 623 DEINKTARDLPFATSYIVIFLYTSLALGYSYSSRVNVDKATL-GLGGVAVVLGAVM 681
Db 338 NEVKTLIE-----AIIIVFLVMYLFQNF-----RATLIPTIAVPVLLGTF 380
Qy 682 AAMGFYSYLGIRSLVILQVFPVLVSVG--ADNIFIVLEYQRLPRRPGEPREVIHGRA 739
Db 381 AILSAFGY-----SINTLTMPG-MVLAIGLLVDDAIVVVENVERVMQBEGLPPKEATKKS 434
Qy 740 LGR-----VAPSMLLCSLSEATCFPLGALTMPAVRTFALTSLGLAVILDFLLQMSAFVAL 794
Db 435 MEQIQGALVGIALLVSAVFPVMAFPGATGAI--YRQFSITIVSAMVLSVLVALILTPAL 492
Qy 795 LSLSKRQEAASLDVCCVCKPOELPPQGGELLLGFFQKAYAPFLHMI-----TR 846
Db 493 CA-----TMLKPKIKGDHGPKTG-FFGWFNNMPEKSTHHYTDSDVANILRST 537
Qy 847 GVVLFLFLALF-----GVSLYSMCHISVGLDQELALPKDSYLLDY 886
Db 538 GRYLVIYLAIVIGMAVLFMRLPSSFLPEDQGVFL-TMVQLPAGATQERTQKVLNHVTDY 596
Qy 887 FL-----FLNRYPEVCAPVYFVTLGYNFSSEAGMN-----AI 919
Db 597 YLDKEKNVNVSVFTVN-----GFGFSQG-QGNTGLAFVSLKNWDERKGEQNKVPAL 646
Qy 920 CSSAG-----CWNFSF-----TKIQYAT 938
Db 647 VSRASAFSKIKDGMVFAFNLPALVELGTATGDFQLIDQGNLGHQQLTDARNQLLGMAA 706
Qy 939 EFPEQSYLAIP-----ASSWDDDFIDWLT 962
Db 707 QHPDMLVGRPNGLGLEDTPQKVEVDQEAQALGVAISDINTTILGSAAGGSYVNDFFD--- 763

Qy 963 PSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQPHKLPWFNLDRPNI 1022
Db 764 RGRVKVYVQADAPFRMLPDDIDK-----WYVRNN--- 793
Qy 1023 KCPKGGGLAAYST-----SVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELA 1073
Db 794 ---MGQMVSPATFSTAKWEYGSPLRYNGLPSPMEILGQAAPGKST---GEAMDLMQELA 847
Qy 1074 ANITADLRKVPCTDPAFEVFPYTITNVFYEQYLT-----ILPGLFMLSLCL----- 1120
Db 848 AKLP SGV-----GYDWTGMSYQERLSGNQAPADYAIASLIVFLCLALYESW 894
Qy 1121 -----VPTFAVSCLLGLDLRLSGLLNLLSTVILVDTVGFPMALWDISYNVSLI-- 1169
Db 895 SIPPSVMLVPLGVVGA-LLAATLR-GLENDVYFQVGLLTIGLSA-----KNAILIVEF 947
Qy 1170 -----NLVSAYGMSVEFVSHITRSPTAISTKPTWLBRAKEATISMGSANVPAG 1215
Db 948 AKDLMDKEGKGLVESTLESVRMLRPILMTSLAPILGVMP-----LVISSG-----AG 995
Qy 1216 VAMTNLPGLVL-CLAKAQLIQIIEFFRLNLLI 1246
Db 996 SGAQNAGVTGVMGGMITATVLAIFVFLFFV 1027

Search completed: April 11, 2006, 00:46:05
Job time : 92 secs

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioceleration Ltd.
 OM protein - protein search, using sw model
 Run on: April 11, 2006, 00:36:46 ; Search time 265 Seconds
 (without alignments)
 3546.278 Million cell updates/sec

Title: US-10-736-769-4
 Perfect score: 6909
 Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGALSNFLPNNGRF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : UniProt 05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6872.5	99.5	1359	1 NPCL1_HUMAN	Q9uhc9 homo sapien
2	5421.5	78.5	1331	1 NPCL1_RAT	Q6t3u3 rattus norv
3	5407	78.3	1333	1 NPCL1_MOUSE	Q6t3u4 mus musculu
4	3179	46.0	1132	2 Q4T749_TETNG	Q4t749 tetraodon n
5	2413	34.9	1277	2 Q9JLG3_CRIGR	Q9jlg3 cricetus
6	2412	34.9	1277	1 NPCL1_FIG	P56941 sus scrofa
7	2405	34.8	1276	2 Q8MKD8_FELCA	Q8mkd8 felis silve
8	2404.5	34.8	1276	2 Q9GK52_CANFA	Q9gk52 canis fami
9	2402.5	34.8	1278	1 NPCL1_HUMAN	O15118 homo sapien
10	2402	34.8	1289	2 Q59GR1_HUMAN	Q59gr1 homo sapien
11	2400	34.7	1277	2 Q7TMD4_MOUSE	Q7tmd4 mus musculu
12	2399	34.7	1276	2 Q9NQOQ_FELCA	Q9nqoq felis silve
13	2395	34.7	1276	2 Q8MI49_FELCA	Q8mi49 felis silve
14	2385	34.5	1278	1 NPCL1_MOUSE	Q35604 mus musculu
15	2367.5	34.3	1277	2 Q9GLC9_BOVIN	Q9glc9 bos taurus
16	2361	34.2	1286	2 Q9TT75_RABIT	Q9tt75 carytolagus
17	2338	33.8	1209	2 Q4RWY5_TETNG	Q4rwy5 tetraodon n
18	2262	32.7	1287	2 Q9VL24_DROME	Q9vl24 drosophila
19	2256	32.7	1287	2 Q9U5W1_DROME	Q9u5w1 drosophila
20	2253	32.6	1287	2 Q7YU59_DROME	Q7yu59 drosophila
21	2205	31.9	1291	2 Q7YU09_ANOGA	Q7y409 anopheles g
22	1864.5	27.0	1233	2 Q9VRC9_DROME	Q9vrc9 drosophila
23	1760	25.5	1003	2 Q7FSO3_ANOGA	Q7ps03 anopheles g
24	1725	25.0	1275	2 Q5LNK7_MAGGR	Q5lnk7 magnaporthe
25	1712.5	24.8	1361	2 Q7XUB7_ORYSA	Q7xub7 oryza sativ
26	1700.5	24.6	1271	2 Q5BBG1_EMENI	Q5bbg1 aspergillus
27	1659.5	24.0	1330	2 Q5KGS9_CRYNE	Q5kgs9 cryptococcu
28	1658.5	24.0	1330	2 Q5SSD4_CRYNE	Q5ssd4 cryptococcu
29	1643	23.8	1273	2 Q4WNG5_ASFFU	Q4wns5 aspergillus
30	1614.5	23.4	1295	2 Q41OK4_GIBBEA	Q41ok4 gibberella
31	1549	22.4	1264	2 Q6BT03_DEBHA	Q6bt03 debaryomyce

ALIGNMENTS

32	1527.5	22.1	1275	2	Q9SHN9_ARATH	Q9shn9 arabidopsis
33	1511.5	21.9	1342	2	Q9TVK6_DICDI	Q9tvk6 dictyosteli
34	1508	21.8	1239	2	Q6CBAL_YARLI	Q6cbal yarrowia li
35	1500	21.7	1162	2	Q7RWL9_NEUCR	Q7rw19 neurospora
36	1461.5	21.2	1256	2	Q592V0_CANAL	Q592v0 candida alb
37	1452.5	21.0	1397	2	Q551C5_DICDI	Q551c5 dictyosteli
38	1437.5	20.8	1489	2	Q4PEB3_USTMA	Q4peb3 ustilago ma
39	1394	20.2	1055	2	Q9SVF0_ARATH	Q9svf0 arabidopsis
40	1386.5	20.1	620	2	Q4H344_CIOIN	Q4h344 ciona intes
41	1359	19.7	1382	2	Q618V4_CAEBR	Q618v4 caenorhabdi
42	1341.5	19.4	1383	1	NPCL1_CAEBL	Q19127 caenorhabdi
43	1329	19.2	1170	2	Q12200_YEAST	Q12200 saccharomyc
44	1323	19.1	1339	2	Q5ORB6_ENTHI	Q5orb6 entamoeba h
45	1312.5	19.0	1178	2	Q750G1_ASHGO	Q750g1 ashbya goss
46	1300.5	18.8	1177	2	Q6CIF2_KLUJA	Q6cif2 kluyveromyc
47	1274.5	18.4	1193	2	Q6FKT2_CANGA	Q6fkt2 candida gla
48	1046	15.1	1274	1	NPCL1_CAEBL	Q19127 caenorhabdi
49	981	14.2	1222	2	Q612M9_CAEBR	Q612m9 caenorhabdi
50	876	12.7	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
51	874	12.7	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
52	873	12.6	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
53	873	12.6	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
54	873	12.6	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
55	869	12.6	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
56	869	12.6	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
57	867	12.5	371	2	Q6GNS9_CIOIN	Q6gns9 ciona intes
58	858.5	12.4	524	2	Q9CF80_ARATH	Q9cf80 arabidopsis
59	854	12.4	371	2	Q6GNS0_CIOIN	Q6gns0 ciona intes
60	776.5	11.2	601	2	Q5TQH9_ANOGA	Q5tqh9 anopheles g
61	651.5	9.4	1447	1	PTCL1_HUMAN	Q13635 homo sapien
62	649	9.4	1434	2	Q6UY90_RAT	Q6uy90 rattus norv
63	646.5	9.4	1434	1	PTCL1_MOUSE	Q61115 mus musculu
64	636.5	9.2	827	2	Q5RJ13_BRARE	Q5rj13 brachydanio
65	628.5	9.1	1442	1	PTCL1_CHICK	Q90693 gallus gall
66	624	9.0	1296	2	Q5VZC2_HUMAN	Q5vzc2 homo sapien
67	617.5	8.9	1239	2	Q9DEF4_XENLA	Q9def4 xenopus lae
68	617.5	8.9	1418	2	Q98SW6_XENLA	Q98sw6 xenopus lae
69	601.5	8.7	1243	2	Q9W6T6_BRARE	Q9w6t6 brachydanio
70	599.5	8.7	1413	2	Q9DEF3_XENLA	Q9def3 xenopus lae
71	599.5	8.7	1422	2	Q6IRA5_XENLA	Q6ira5 xenopus lae
72	598.5	8.7	933	2	P91129_CAEBL	P91129 caenorhabdi
73	597	8.6	896	2	Q617L5_CAEBR	Q617l5 caenorhabdi
74	592	8.6	900	2	P91346_CAEBL	P91346 caenorhabdi
75	589	8.5	278	2	Q4H343_CIOIN	Q4h343 ciona intes
76	585	8.5	933	2	Q61EN1_CAEBR	Q61en1 caenorhabdi
77	584	8.5	1203	1	PTC2_HUMAN	Q9V6C5 homo sapien
78	584	8.5	1203	2	Q53Z57_HUMAN	Q53z57 homo sapien
79	583.5	8.4	1182	1	PTC2_MOUSE	Q35395 mus musculu
80	573.5	8.3	1003	2	Q9XXR9_CAEBL	Q9xxr9 caenorhabdi
81	572	8.3	1220	1	PTCL1_BRARE	Q98864 brachydanio
82	566.5	8.2	1004	2	Q60VB2_CAEBR	Q60vb2 caenorhabdi
83	562.5	8.1	1169	2	Q9V9L4_DROME	Q9v9l4 drosophila
84	532	7.7	816	2	Q4SP00_TETNG	Q4sp00 tetraodon n
85	522.5	7.6	1318	2	Q9XYP5_JUNCO	Q9xyp5 junonia coe
86	522	7.5	1197	2	Q61Z32_CAEBR	Q61z32 caenorhabdi
87	521.5	7.5	1197	2	Q4RVD6_TETNG	Q4rvd6 tetraodon n
88	514	7.4	1405	1	PTCL1_CAEBL	Q96114 caenorhabdi
89	474	6.9	1286	1	PTC1_DROME	P18502 drosophila
90	473.5	6.9	1015	2	Q99338_CAEBL	Q99338 caenorhabdi
91	470.5	6.8	915	2	Q61PC1_CAEBR	Q61pc1 caenorhabdi
92	464	6.7	974	2	Q621T1_CAEBR	Q621t1 caenorhabdi
93	463	6.7	1380	2	Q60VB7_CAEBR	Q60vb7 caenorhabdi
94	458.5	6.6	955	2	Q19945_CAEBL	Q19945 caenorhabdi
95	455	6.6	1358	2	Q6AW15_CAEBL	Q6aw15 caenorhabdi
96	453.5	6.6	1361	2	Q6AW16_CAEBL	Q6aw16 caenorhabdi
97	444.5	6.4	844	1	PT89_CAEBL	Q03602 caenorhabdi
98	436	6.3	956	2	Q61829_CAEBL	Q61829 caenorhabdi
99	424	6.1	951	2	Q60W90_CAEBR	Q60w90 caenorhabdi
100	422.5	6.1	939	2	Q7Q2Y4_ANOGA	Q7q2y4 anopheles g

RESULT 1
ID NPC1L1 HUMAN STANDARD; PRT; 1359 AA.
AC Q9UHC9; Q6R304; Q9UHC9;
DT 13-SEP-2005 (Rel. 48, Created)
DT 13-SEP-2005 (Rel. 48, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Niemann-Pick C1-like protein 1 precursor.
GN Name=NPC1L1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA] (ISOFORMS 1; 2 AND 3), AND TISSUE
RP SPECIFICITY.
RX MEDLINE=20247253; PubMed=10783261; DOI=10.1006/geno.2000.6151;
RA Davies J.P., Levy B., Ioannou Y.A.;
RT Evidence for a Niemann-Pick C (NPC) gene family: identification and
RT characterization of NPC1L1.";
RL Genomics 65:137-145(2000).
[2]
RN NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 2), AND TISSUE SPECIFICITY.
RX PubMed=14976318; DOI=10.1126/science.1093131;
RA Altman S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
RA Tetzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
RA Murgolo N., Graziano M.P.;
RT "Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol
RT absorption.";
RL Science 303:1201-1204(2004).
[3]
RN TISSUE SPECIFICITY, AND SUBCELLULAR LOCATION.
RX PubMed=15671032; DOI=10.1074/jbc.M409110200;
RA Davies J.P., Scott C., Oishi K., Liapis A., Ioannou Y.A.;
RT "Inactivation of NPC1L1 causes multiple lipid transport defects and
RT protects against diet-induced hypercholesterolemia.";
RL J. Biol. Chem. 280:12710-12720(2005).
[4]
RN INDUCTION.
RX PubMed=15604518; DOI=10.1194/jlr.M400400-JLR200;
RA van der Veen J.N., Kruit J.K., Havings R., Baller J.F.W., Chimini G.,
RA Lestavel S., Staelens B., Groot P.H.E., Groen A.K., Kuipers F.;
RT "Reduced cholesterol absorption upon PPARdelta activation coincides
RT with decreased intestinal expression of NPC1L1.";
RL J. Lipid Res. 46:526-534(2005).
[5]
RN FUNCTION.
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
RA Garcia-Calvo M., Lisnock J., Bull H.G., Hawes B.E., Burnett D.A.,
RA Braun M.P., Crona J.H., Davis J.H., Jr., Dean D.C., DeWerns P.A.,
RA Graziano M.P., Hughes M., Macintyre D.E., Ogawa A., O'Neill K.A.,
RA Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
RA Ujainwalla F., Altman S.W., Chapman K.T., Thornberry N.A.;
RT "The target of ezetimibe is Niemann-Pick C1-like 1 (NPC1L1).";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137(2005).
[6]
RN VARIANTS LEU-55 AND ASN-1233.
RX PubMed=15679830; DOI=10.1111/j.1399-0004.2004.00388.x;
RA Wang J., Williams C.M., Hegele R.A.;
RT "Compound heterozygosity for two non-synonymous polymorphisms in
RT NPC1L1 in a non-responder to ezetimibe.";
RL Clin. Genet. 67:175-177(2005).
[7]
RN FUNCTION: Play a major role in cholesterol homeostasis. Is
RN critical for the uptake of cholesterol across the plasma membrane
RN of the intestinal enterocyte. Is the direct molecular target of
RN ezetimibe, a drug that inhibits cholesterol absorption. Lack of
RN activity leads to multiple lipid transport defects. The protein
RN may have a function in the transport of multiple lipids and their
RN homeostasis, and may play a critical role in regulating lipid
RN metabolism.
RN -1- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
of brush border membranes from proximal enterocytes suggests
considerable association with the apical membrane fraction. Exists
as a predominantly cell surface membrane expressed protein (By
similarity). According to Ref.3 localizes in a subcellular
vesicular compartment rich in RAB5.
-1- ALTERNATIVE PRODUCTS:
Name=1;
IsoId=Q9UHC9-1; Sequence=Displayed;
Name=2; Synonyms=NPC1L1DELTAEL5;
IsoId=Q9UHC9-2; Sequence=VSP_015314;
Name=3; Synonyms=NPC1L1;
IsoId=Q9UHC9-3; Sequence=VSP_015312, VSP_015313;
TISSUE SPECIFICITY: Widely expressed. Expressed in liver. Also
expressed in small intestine, pancreas, kidney, lung, pancreas,
spleen, heart, gall bladder, brain, testis, stomach and muscle.
-1- INDUCTION: Expression is decreased in Caco-2 cells upon PPARd
activation.
-1- PTM: Highly glycosylated (By similarity).
-1- POLYMORPHISM: Variations in NPC1L1 gene could be associated with
nonresponse to ezetimibe treatment
-1- SIMILARITY: Belongs to the patched family.
-1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
EMBL; AF192522; AAF20396.1; -; mRNA.
EMBL; AF192523; AAP20397.1; -; mRNA.
EMBL; AY515256; AAS56939.1; -; mRNA.
EMBL; AY437885; AAR57886.1; -; mRNA.
DR Ensembl; ENSG0000015520; Homo sapiens.
HGNC; HGNC:7898; NPC1L1.
MIM; 608010; --
GO; GO:0015918; P:sterol transport; TAS.
InterPro; IPR004765; NP_C type.
InterPro; IPR003392; Patched.
InterPro; IPR000731; SSD_5TM.
Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
PROSITE; PS50156; SSD; 1.
KW Alternative splicing; Cholesterol metabolism; Glycoprotein;
KW Lipid metabolism; Polymorphism; Signal; Steroid metabolism;
KW Transmembrane.
FT SIGNAL 1 21 Potential.
FT CHAIN 1 21 Niemann-Pick C1-like protein 1.
FT TOPO_DOM 22 284 Extracellular (Potential).
FT TRANSMEM 285 305 1 (Potential).
FT TOPO_DOM 306 351 Cytoplasmic (Potential).
FT TRANSMEM 352 372 2 (Potential).
FT TOPO_DOM 373 632 Extracellular (Potential).
FT TRANSMEM 633 653 3 (Potential).
FT TOPO_DOM 654 666 Cytoplasmic (Potential).
FT TRANSMEM 667 687 4 (Potential).
FT TOPO_DOM 688 696 Extracellular (Potential).
FT TRANSMEM 697 717 5 (Potential).
FT TOPO_DOM 718 742 Cytoplasmic (Potential).
FT TRANSMEM 743 763 6 (Potential).
FT TOPO_DOM 764 776 Extracellular (Potential).
FT TRANSMEM 777 797 7 (Potential).
FT TOPO_DOM 798 846 Cytoplasmic (Potential).
FT TRANSMEM 847 867 8 (Potential).
FT TOPO_DOM 868 882 Extracellular (Potential).
FT TRANSMEM 883 903 9 (Potential).
FT TOPO_DOM 904 1139 Cytoplasmic (Potential).
FT TRANSMEM 1140 1160 10 (Potential).
FT TOPO_DOM 1161 1168 Extracellular (Potential).
FT TRANSMEM 1169 1189 11 (Potential).
FT TOPO_DOM 1190 1191 Cytoplasmic (Potential).
FT TRANSMEM 1192 1212 12 (Potential).

FT	TOPO_DOM	1213	1236	Extracellular (Potential).
FT	TRANSMEM	1237	1257	13 (Potential).
FT	TOPO_DOM	1258	1359	Cytoplasmic (Potential).
FT	DOMAIN	632	797	SSD.
FT	COMBIAS	500	503	Poly-Leu.
FT	CARBOHYD	54	54	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	132	132	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	138	138	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	244	244	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	416	416	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	431	431	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	464	464	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	479	479	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	497	497	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	506	506	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	626	626	N-linked (GlcNAc. .) (Potential).
FT	VARSPLIC	723	724	RL -> GP (in isoform 3).
FT				/FTid=VSP 015312.
FT	VARSPLIC	725	1359	Missing (in isoform 3).
FT				/FTid=VSP 015313.
FT	VARSPLIC	1046	1072	Missing (in isoform 2).
FT				/FTid=VSP 015314.
FT	VARIANT	55	55	V -> L (nonresponse to ezetimibe treatment).
FT				/FTid=VAR 023369.
FT	VARIANT	1233	1233	I -> N (nonresponse to ezetimibe treatment).
FT				/FTid=VAR 023370.
SQ	SEQUENCE	1359 AA; 148698 MW; 3225D5D93B768B6 CRC64;		
Query Match 99.5%; Score 6872.5; DB 1; Length 1359;				
Best Local Similarity 97.9%; Pred. No. 0;				
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;				
QY	1	MAEAGLRGHLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNSCLSN	60	
DB	1	MAEAGLRGHLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNSCLSN	60	
QY	61	TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVSEASLSITKALLTRCPACSDNF	120	
DB	61	TPARKITGDHLILLOKICPRLYTGPNTOACCSAKOLVSEASLSITKALLTRCPACSDNF	120	
QY	121	VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180	
DB	121	VNLHCHTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA	180	
QY	181	ATLAVGTMCVYGSALCNQRLNFGDGTGNGLAFLDITPHLLFPGQAVGSGIQPLNEGV	240	
DB	181	ATLAVGTMCVYGSALCNQRLNFGDGTGNGLAFLDITPHLLFPGQAVGSGIQPLNEGV	240	
QY	241	ARCNESQGDVATCSQDCCAACPAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI	300	
DB	241	ARCNESQGDVATCSQDCCAACPAARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI	300	
QY	301	LIVGFRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWELTILVLSV	360	
DB	301	LIVGFRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQFPQGWGTWVASWELTILVLSV	360	
QY	361	IPVVALAAGLVTELTDPVELWSPNSQARSEKAFHDQHPFRFTNQVILTAENRSY	420	
DB	361	IPVVALAAGLVTELTDPVELWSPNSQARSEKAFHDQHPFRFTNQVILTAENRSY	420	
QY	421	RYDSLILGPKNFGSLDLDLLELEQLERLRLQVMSPEAORNI SLQDICVAPLNPDNT	480	
DB	421	RYDSLILGPKNFGSLDLDLLELEQLERLRLQVMSPEAORNI SLQDICVAPLNPDNT	480	
QY	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKXDHFLYCANAPLTKDGTALAL	540	
DB	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKXDHFLYCANAPLTKDGTALAL	540	
QY	541	SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEBM	600	
DB	541	SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEBM	600	

QY	601	RAFORMAGMFQVTTFAERSLEDEINRTTAEDLPATSYIVIFLYISIALGSSWSRV	660	
DB	601	RAFORMAGMFQVTTFAERSLEDEINRTTAEDLPATSYIVIFLYISIALGSSWSRV	660	
QY	661	MYDSKATILGLGVAVVILGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGDANIFLVL	720	
DB	661	MYDSKATILGLGVAVVILGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGDANIFLVL	720	
QY	721	YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGLTTPMPAVRTTALTSGLAV	780	
DB	721	YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGLTTPMPAVRTTALTSGLAV	780	
QY	781	ILDPLLOMSAFVALLSLDSKRQASRLDVCCCKVQBELPPPGQGGEGLLGFFQKAYAPFL	840	
DB	781	ILDPLLOMSAFVALLSLDSKRQASRLDVCCCKVQBELPPPGQGGEGLLGFFQKAYAPFL	840	
QY	841	LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900	
DB	841	LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900	
QY	901	YFVTTGLGNFSSEAGMNAICSSAGCNFNSFTOKIOVATFPFQSYLAIPASSWDDFDIOW	960	
DB	901	YFVTTGLGNFSSEAGMNAICSSAGCNFNSFTOKIOVATFPFQSYLAIPASSWDDFDIOW	960	
QY	961	LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSTITMGSVRPSVEQFHKYLFPWFLNDRP	1020	
DB	961	LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSTITMGSVRPSVEQFHKYLFPWFLNDRP	1020	
QY	1021	NIKCPKGGGLAAYSTVNLTSDQVLTVAISPRLEYSCTISAHGNLYLLDSASRPMAYH	1080	
DB	1021	NIKCPKGGGLAAYSTVNLTSDQVLTVAISPRLEYSCTISAHGNLYLLDSASRPMAYH	1080	
QY	1054	KPKNSQDYTEALRAARELAANITADLRKVPCTDPAPFPVPTITNVFYEQVLTILPEGL	1113	
DB	1081	KPKNSQDYTEALRAARELAANITADLRKVPCTDPAPFPVPTITNVFYEQVLTILPEGL	1140	
QY	1114	FMLSICLVPTFAVSCLLGLDLRLSGLNLLSVMLIVDTVGFMAWLDISYNAVSLINLVS	1173	
DB	1141	FMLSICLVPTFAVSCLLGLDLRLSGLNLLSVMLIVDTVGFMAWLDISYNAVSLINLVS	1200	
QY	1174	AVGMSVFEVSHITRFAISTKPTWLERAKBATISMGSAVAGVAMTNLPGILVLGLAKAQ	1233	
DB	1201	AVGMSVFEVSHITRFAISTKPTWLERAKBATISMGSAVAGVAMTNLPGILVLGLAKAQ	1260	
QY	1234	LQIIEFFRLNLLITLLGLLHGLVFLPVILSYVGPVDPNPALEOKRAEAAVAVWVASC	1293	
DB	1261	LQIIEFFRLNLLITLLGLLHGLVFLPVILSYVGPVDPNPALEOKRAEAAVAVWVASC	1320	
QY	1294	NHPSRVSTADNIYVNHSPSGSIKGAGAINFLPNNGROF	1332	
DB	1321	NHPSRVSTADNIYVNHSPSGSIKGAGAINFLPNNGROF	1359	

RESULT 2

NPCL1 RAT	STANDARD;	PRT;	1331 AA.
ID NPCL1 RAT			
AC O6T3U3			
DT 13-SEP-2005 (Rel. 48, Created)			
DT 13-SEP-2005 (Rel. 48, Last sequence update)			
DT 13-SEP-2005 (Rel. 48, Last annotation update)			
DE Niemann-Pick C1-like protein 1 precursor.			
GN Name=Npc1l1;			
OS Rattus norvegicus (Rat).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;			
OC Murioidea; Muridae; Murinae; Rattus.			
OX NCBI TaxID=10116;			
RN (1)			
RP NUCLEOTIDE SEQUENCE [MRNA], TISSUE SPECIFICITY, AND SUBCELLULAR			
RC LOCATION.			
RC STRAIN=Sprague-Dawley;			
RX PubMed=14976318; DOI=10.1126/science.1093131;			

RA Altmann S.W., Davis H.R. Jr., Zhu L.-J., Yao X., Hoos L.M.,
RA Terzloff G., Iyer S.P.N., Maguire M., Golovko A., Zeng M., Wang L.,
RA Murgolo N., Graziano M.P.;
RT "Niemann-Pick C1 like 1 protein is critical for intestinal cholesterol
RL absorption.";
RN Science 303:1201-1204 (2004).
RP [2]
RP SUBCELLULAR LOCATION, AND CHARACTERIZATION.
RX PubMed=15928087; DOI=10.1073/pnas.0500269102;
RX Garcia-Calvo M., Lissack J., Bull H.G., Hawes B.E., Burnett D.A.,
RX Braun M.P., Crona J.H., Davis H.R. Jr., Dean D.C., Detmers P.A.,
RA Graziano M.P., Hughes M., Macintyre D.E., Ogawa A., O'Neill K.A.,
RA Iyer S.P.N., Shevell D.E., Smith M.M., Tang Y.S., Makarewicz A.M.,
RA Ujjainwalla F., Altmann S.W., Chapman K.T., Thornberry N.A.;
RT "The target of ezetimibe is Niemann-Pick C1-Like 1 (NPC1L1).";
RL Proc. Natl. Acad. Sci. U.S.A. 102:8132-8137 (2005).
CC -!- FUNCTION: play a major role in cholesterol homeostasis. Is
CC critical for the uptake of cholesterol across the plasma membrane
CC of the intestinal enterocyte. Is the direct molecular target of
CC ezetimibe, a drug that inhibits cholesterol absorption (by
CC similarity). Lack of activity leads to multiple lipid transport
CC defects. The protein may have a function in the transport of
CC multiple lipids and their homeostasis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Subfractionation
CC of brush border membranes from proximal enterocytes suggests
CC considerable association with the apical membrane fraction. Exists
CC as a predominantly cell surface membrane expressed protein.
CC -!- TISSUE SPECIFICITY: Small intestine showed the highest level of
CC expression. Expression in other tissue including gall bladder,
CC liver, testis and stomach is also observed. Along the duodenum-
CC ileum axis, the levels vary in different segments of intestine,
CC with peak expression in the proximal jejunum. Protein expression
CC is confined to the enterocyte. Discrete localization to the
CC epithelial layer bordering the luminal space along the crypt-
CC villus axis. Protein expression in the enterocyte is observed
CC closest to the luminal space. Expression in enterocytes from the
CC proximal (jejunum) but not in the distal (ileum) region.
CC -!- PTM: Highly glycosylated.
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AY437867; AAR97888.1; -; mRNA.
CC RGD: 1303135; Npc1l1.
CC InterPro: IPR004765; NP_C type.
CC InterPro: IPR003392; Patched.
CC InterPro: IPR000731; SSD_5TM.
CC Pfam: PF02460; Patched, 1.
CC TIGRFAMs: TIGR00917; 2A060601; 1.
CC PROSITE: PS50156; SSD, 1.
CC Cholesterol metabolism; Glycoprotein; Lipid metabolism; Signal;
KW Steroid metabolism; Transmembrane.
FT SIGNAL 1 20 Potential.
FT CHAIN 21 1331
FT TOPO_DOM 21 282 Extracellular (Potential).
FT TRANSMEM 283 303 1 (Potential).
FT TOPO_DOM 304 352 Cytoplasmic (Potential).
FT TRANSMEM 353 373 2 (Potential).
FT TOPO_DOM 374 632 Extracellular (Potential).
FT TRANSMEM 633 653 3 (Potential).

FT TOPO_DOM 654 665
FT TRANSMEM 666 686
FT TOPO_DOM 687 696
FT TRANSMEM 697 717
FT TOPO_DOM 718 742
FT TRANSMEM 743 763
FT TOPO_DOM 764 776
FT TRANSMEM 777 797
FT TOPO_DOM 798 846
FT TRANSMEM 847 867
FT TOPO_DOM 868 1113
FT TRANSMEM 1114 1134
FT TOPO_DOM 1135 1142
FT TRANSMEM 1143 1163
FT TOPO_DOM 1164 1165
FT TRANSMEM 1166 1186
FT TOPO_DOM 1187 1206
FT TRANSMEM 1207 1227
FT TOPO_DOM 1228 1242
FT TRANSMEM 1243 1263
FT TOPO_DOM 1264 1331
FT DOMAIN 632 797
FT CARBOHYD 53 53
FT CARBOHYD 85 85
FT CARBOHYD 138 138
FT CARBOHYD 244 244
FT CARBOHYD 416 416
FT CARBOHYD 431 431
FT CARBOHYD 464 464
FT CARBOHYD 479 479
FT CARBOHYD 497 497
FT CARBOHYD 506 506
FT CARBOHYD 606 606
FT CARBOHYD 626 626
FT CARBOHYD 909 909
FT CARBOHYD 917 917
FT CARBOHYD 996 996
FT CARBOHYD 1038 1038
FT CARBOHYD 1076 1076
SQ SEQUENCE 1331 AA; 146415 MW; 2E10EF2E3A337F70 CRC64;
Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
QY 1 MAEAGLRCGLWALLLLRLAQSEPVYTHIQPGYCAFYDECGKNPELSGLSLTSLNVCSLSN 60
DB 1 MAEAWL-GWLLWALLLLSAAQGLYTPKHEAGVCTFYECGKNPELSGLSLTSLNVCSLSN 59
QY 61 TPARKITGDHLILQKICPRLYTGPNTQ-ACCSAKQLVSLBASISITKALLTRCPACSDN 119
DB 60 TPARRHTGHEHLALLQRICPRLYNGENTTFACCTKQLLSLESSMSITKALLTRCPACSDN 119
QY 120 FVNLCHNTCSNQLSFLINVTVAQLGAGQIPAVVAEAFYQHSFAEQSYDSCSRVPA 179
DB 120 FVSLCHNTCSNQLSFLINVTVAQLGAGQIPAVVAEAFYQHSFAEQSYDSCSRVPA 179
QY 180 AATLAVGTMCVGYSGALCNQRLNFGQDTGNGLAPLDITFHLLPFGQAVSGIQLNEG 239
DB 180 AASLAVGSMGVGYSGALCNQRLNFGQDTGNGLAPLDITFHLLPFGQAVSGIQLNEG 239
QY 240 VARNCSQGDVATCSCDCAASCPAIARPOALDSTFYLGOMPGSLVLIILCSFPAVT 299
DB 240 IAPCNESQGDVATCSCDCAASCPVIPPPEALRPSFYMGPMGWLIIITFVAVFLIS 299
QY 300 ILLVCFRYPAPADKSKMDVPPKGTSLSDKLSFSTHTLLGQFGQGTWVASWPLTILVLS 359
DB 300 AVLVELRVVNRNKNKAEQPOEAPKPKHKLSPHTILGRFQNGTRVVASWPLTILVLS 359
QY 360 VTPVVALAAGLVFTLTTPVLMWSPNSQARSEKAFHDQHFPGPFRTNQVILTPANRSS 419
DB 360 FIVVIALAAGLVFTLTTPVLMWSPNSQARSEKAFHDQHFPGPFRTNQVILTPANRSS 419

DR GO: 0008158; F.hedgehog receptor activity; IEA.

DR GO: 0003001; P.cholesterol transporter; IEA.

DR InterPro; IPR004765; NP_C type.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR000731; SSD 5TM.

DR Pfam; PF02460; Patched; 1.

DR TIGRFAMs; TIGR00917; 2A060601; 1.

DR PROSITE; P050156; SSD; 1.

SQ SEQUENCE 1276 AA; 141695 MW; 07D1327B51BD5681 CRC64;

Query Match 34.8%; Score 2404.5; DB 2; Length 1276;

Best Local Similarity 40.4%; Pred. No. 4.2e-161;

Matches 524; Conservative 233; Mismatches 434; Indels 105; Gaps 28;

QY 33 CAFYDECGRNPELSGLMTLSNVSCLSNTPARKITGDHLLILQKICPRLYTGENTQACCS 92

DB 25 CVWYGECS--IASGD---KRYNCQYSGPPKPLPKDGYDLAQELCPGLFF-DNVSVCDD 76

QY 93 AKQLVSLASLITKALLTRCPACSDNFVNLHCHNCTSPNQLFINVTR-----VAQLGAG 148

DB 77 VQQLRLTKDLSQLPLQFLSRCPSCFYNLNLFCELTCSPRQSQFLNVTETEDYDVPVTNQ 136

QY 149 QLPVAVYAEAFQHGSPAEQSDYSCSRVRVPAATLAVGTMCVGYSGALCNAQRLNFGQD 208

DB 137 TKTNVKELQYVCGESFANAMYNACRDVEAPSSNDKALGLLCKGAEAE-CNATNWIIEYFN 195

QY 209 TGNGLAPL-----DITPHLLEPGQAVSGIGIQLNEGVARCNESQGDVATCSQDCCAA 261

DB 196 KONGQAPFTIIFSDLPALH-----GKMPNNATKGCDEPDEVETAPCSQDCSV 245

QY 262 SC-----PAIRAPQALDSTFYLGQMPGSLVLIILCSVFVAVVITLLVGRVAPARD 312

DB 246 VCGPKPQPPAPAPARILGLDAMVIMWTYMAFLMPPGAFPAVWCY----- 293

QY 313 KSKWDPKGTSLSDKLSSTHT-----LLGQPFQ-----WGTWVASWPLT 354

DB 294 -RRKYFVSEYTPIDNSIAFVNAGDTGEASCCDALGAFFEGCLRLFTQWGSCFIRNPGC 352

QY 355 ILVLSVIPVVALAAGLVFELTTDPVELMSAPNSQARSEKAFDQHFPGFPFTNNOVILTA 414

DB 353 IIFSLAFTACSGLVFSEVTVNTPVDLMSAPGQARLEKYFDHAFGFPFTEQLIIQA 412

QY 415 PNRSYRYDSSLGPK-NPFGILDLDLLELELEQLERHLQWGPAPQARNISLQDICVA 473

DB 413 PHTSVHTYQYPSPGSDVPPGPDLDIGILHVLQDLQTAIEN--ITATYNNETVTLQDICVA 470

QY 474 PLNPNTSYLDCINSILLOYFQNNRTLLLTANTQILMGTSQV--DWKDHFLYCANAPLT 531

DB 471 PLSPYNNK---NCTIMSVLNYFQNSHML-----DHKIGDDFYVYADYHTHLLYCVRAPAS 522

QY 532 PKDGTALALSCMADYCAPVFPFLAIGYKGDYSBAELIMTFSLNNYPAGDPRLAQAKL 591

DB 523 LNDTSLLHDPCLGTGGPVPFPMVLVGGYDDQNNATALVITFPVNNYNTDEKQRAQA 582

QY 592 WEEAFLEENRAFQRRMAGMFQVFTFAERSLEDEINRTAEDLPFATSYVIFLYISLAL 651

DB 583 WEKEFINFKVYEN---PNLTISFTTERSIEDLNRESNGDFTVLISYAVMFLYISIAL 639

QY 652 GYSYNSRVMVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVPLVLSVGA 711

DB 640 GHKSCSRLVDSKISLTAGILVILSVSCSLGIFSYFGIPLTIVIEVIPFLVAVGV 699

QY 712 DNIETVLEYQRLPRPGEPREVIHGRALGRVAPSMILCSLSEACFFLGALTPPAVPT 771

DB 700 DNIETVLEYQRLPRPGEPREVIHGRALGRVAPSMILCSLSEACFFLGALTPPAVPT 759

QY 772 FALTGLAVILDQLMSAFVALLSLDSKQBSRLDVCCVKPQELPPGQG-----EGL 827

DB 760 FSLPAGMAVLIDFLQITCFVSLGLDLKQENRLDLCLLTGSE---GGTGIQASESC 816

QY 828 LIGFTQKAYAPFLHWTIGVVLVLLFLALFGVSLYSMSCHISVGLDQELAPKDSYLLDYF 887

DB 817 LPRFFKNSYSPFLKDWMPRIVIAFVFGILSPSLAVLNKVKETGLDQSLSPDSDYMDYF 876

QY 888 LFLNRYFEVGAPVYFVTTLTGYNFSEAGMNAICSSAGCNFNFTQKIQIYATEPPEQSYLA 947

DB 877 KSL-KYLHAGPPVYFVLEBEGHDYTSLEGQNMVCGGMCNDSLVQIQPSAAQLDNYTHIG 935

QY 948 IPASSWDDFDWLTP-SSCCLYISGPNKDFCFSTVNSLCLNCKMCSITM-GSVRPSV 1005

DB 936 PAPSSWIDDDYDFWVKPQSSCCRVNS---TDQFCNASVVDPRACVR-CRELTQEGKRRPQ 991

QY 1006 EGFHYKLPWFNDNRPNIKCPKGLAAYSTSVNLT-SDGOVLASRFMAYHKPKNSQDYTE 1064

DB 992 EDFAPFLPMFLSDNPNPKCKGCHAYGSAVNLVGNDSVSGATYFWYTHVTVLTQTSADFTD 1051

QY 1065 ALRAARELAANITADLRKVPGTDPAPFVPPYITNMFYEQYLTILPEGLFMLSCLVPTTF 1124

DB 1052 AMRKAILIASNTKTM-SLKGSN--YRVFPYFVYFVYEQYLTITDDTFNLSVSLGALF 1108

QY 1125 AVSCLLGLDLBSGLNLLSIYMLVDTVGFALMDISNVLNLSVAVGMSVBFVSH 1184

DB 1109 LVTVLGCELMSAVIMCVTIAMILVNMFGVNMGLGISLNVSLVNLVMSCGISVBFCSH 1168

QY 1185 ITRSPAISTKPTMLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQLIOLFFPRML 1244

DB 1169 ITRATVSAKGRVERAREALSHMGSSVFSGITLTKFGIVVLAFAKSOIQFOVFTITGMYL 1228

QY 1245 LITLGLLHGLVFLVILSYGVDVNPALAL-EQKR 1279

DB 1229 AMVLGATHGLIFLPLVLLSYIGPSINKAKSLASQER 1264

RESULT 9

NPCL HUMAN

ID NPCL HUMAN STANDARD; PRT; 1278 AA.

AC O15118; Q9P130;

DT 30-MAY-2000 (Rel. 39, Created)

DT 10-MAY-2005 (Rel. 47, Last sequence update)

DT 10-MAY-2005 (Rel. 47, Last annotation update)

DE Niemann-Pick C1 protein precursor.

GN Name=NPCL;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;

OC Homo.

OX NCBI_TaxID=9606;

RN (1)

RP NUCLEOTIDE SEQUENCE, VARIANT ILE-642, AND VARIANTS NPCL

RX MEDLINE=97362323; PubMed=9211849; DOI=10.1126/science.277.5323.228;

RA Carstee E.D., Morris J.A., Coleman K.G., Loftus S.K., Zhang D.,

RA Cummings C., Gu J., Rosenfeld M.A., Favan W.J., Krizman D.B.,

RA Nagle J., Polymeropoulos M.H., Sturley S.L., Ioannou Y.A.,

RA Higgins M.E., Conly M., Cooney A., Brown A., Kaneski C.R.,

RA Blanchette-Mackie E.J., Dwyer N.K., Neufeld E.B., Chang T.-Y.,

RA Liscum L., Strauss J.F. III, Olmo K., Zeigler M., Carmi R., Sokol J.,

RA Markie D., O'Neill R.P., van Diggelen O.P., Ellender M.,

RA Patterson M.C., Brady R.O., Vanier M.T., Pentchev P.G., Tagle D.A.;

RT "Niemann-Pick C1 disease gene: homology to mediators of cholesterol

RT homeostasis."

RL Science 277:228-231 (1997).

RN (2)

RP NUCLEOTIDE SEQUENCE, AND VARIANTS.

RX MEDLINE=99355599; PubMed=10425213; DOI=10.1006/bbrc.1999.1070;

RA Morris J.A., Zhang D., Coleman K.G., Nagle J., Pentchev P.G.,

RA Carstee E.D.;

RT "The genomic organization and polymorphism analysis of the human

RT Niemann-Pick C1 gene."

RL Biochem. Biophys. Res. Commun. 261:493-498 (1999).

RN (3)

RP NUCLEOTIDE SEQUENCE, AND VARIANT ILE-642.

RX MEDLINE=21623216; PubMed=11754101; DOI=10.1002/humu.10016;

RA Bauer P., Kuestermann R., Bauer C., Finckh U., Hufen A., Kropp J.,

RA Braun S., Kuestermann-Kuhn B., Schmidt D., Harzer K., Rolfs J.,

RT "NPCL: complete genomic sequence, mutation analysis, and

RT characterization of haplotypes."


```
QY 164 PAESYDSCSRVRPAAATLAVGTWCGYGSALCNAQWLNLFQDGTGGLAPLDT 219
DB 163 FANAMYACRDEVEAPSNKALGLLGGKADA-CNATNMIBYMFNKONGQAPFTITPVFS 221
QY 220 -PHLLEPQAVSGIOPLNEGVARNESQGDVATCSQDCAAACPAIARPO----- 270
DB 222 DFPV-----RGMEPNNAATKGDSESVDEVTAFCQDCDSIVCGPKPQPPPPAPWTI 273
QY 271 -ALDSTFYLGMPGSLVLIILCSVFAVVITLL-----VGFRVAPARDKSKM 316
DB 274 LGLDAMYIMWTYMAELVFFGAPFAVVCYKRVFVSEYTPIDSNIAFSV-NASDKGE- 331
QY 317 VDPKGTSLSKLSSTHTLQGFQFGWGTWASWPLTILVSLVPIVVALAAGLVFTLT 376
DB 332 -----ASCCDFVSAAPFRCGLRLRTRWGSFCVRNPGCVIIFSLVITACSSGLVFRVT 385
QY 377 TDPVELWAPNSQARSEKAFHDQHFPPFRNTQVILTAPNRSSVRYDSLLGPK-NFSGI 435
DB 386 TNPVDLWAPSQAARLEKEYFDQHFPPFRTEQLIRAPLTDKHIYQPPSGADVPPGPP 445
QY 436 LDLDLLELLELQERLRLQWSPQAQNI SLQDICYAPLNPDPNTSLYDCINSLLQVFO 495
DB 446 LDIQILHQLVDLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT---NCTLSVLYNFO 500
QY 496 NNRITLLLTANQTLMGQTSQDWDKDFELCANAPLTPKDGTAALSCWADYGAPVPEPLA 555
DB 501 NSHSLVDHKKGDFFF---VYADYTHFELYCVNRPASLNDTSLHDPCLGTGGPVFPLV 557
QY 556 IGGYKGYSEALIMTFLSNYPAGDPPRLAQALWBEAFLEENRAPQRRMAGMFQVTF 615
DB 558 LGYDDQYNNATLVIPFVNNYNDIEKQRAQWEKEFINVKYKN---PNLTISF 614
QY 616 TAERLEDEINTTADLPFATSVIPIFYLSLALGYSVSNRVMWDSKATLGLGVAV 675
DB 615 TAERSIEDELNESDSDVFTVSVIYALMELYLSLALGHMKSCRRLLVDSKSLGIAGILI 674
QY 676 VLGAVMANGFSGYLSIRSVILQVPELVLSVGDADNIFIVLSEYQRLPRPGRPREVH 735
DB 675 VLSSVACSGVFSYGLPLTLVIEVIFPLVAVGVDNIFILVQAYQDERLQGTTLQOQ 734
QY 736 IGRALGRVAPSMLLCSLSBAICFFLGPALTPMPAVRTFALTSLGLVILDFLLQMSAFVALL 795
DB 735 LGRVLGEVAPSMFLSFSFSTVAFGLGALSVMVAHTFSLFAGLAVFIDFLQITCFVSL 794
QY 796 SLDSKQERASRLDVCCKVQELPPQO-GEGLLGFQKAYAPFLHWHITRGVLLFL 854
DB 795 GLDIKQEKRLDIFCCVGAEDGTSVQASESCLFRFFKNSYSLPLLKDMRPIVIAIFV 854
QY 855 ALPGVSLYNSCHTSVGLDQELALPKDSYLLDYFLNRYFEVGAPEVFTTLGYNFSSEA 914
DB 855 GVLSFSLVNLKVDIGLQSLSPDDSDSYWVDYFKSISQYLHAGPPYFVLEGHDTSK 914
QY 915 GMAICSSAGCNFFSTQIQVATEPPEOSYLAIPASSWVDVDFIDMLTP-SSCCRYISG 973
DB 915 QGNWVCGGNCNDSLVIQIFNAALQDNYTRIGFAPSSMIDDYFDWVWQSSCCRV---D 971
QY 974 PTKDPCPSTVNSLNCNKMSIT-NGSVRPSVEQHKYLPFLNDRPNIKCPKGLAAY 1032
DB 972 NITDQFCNASVDPACVR-CRPLTPBQKRPQGGDFMRFLPFLSDNPNPKCGKGHAAY 1030
QY 1033 STSVN--LTSQGOVLASRFWAYHKPLKNSQDYTEALRAARELANITADLRKVPVGDPAF 1090
DB 1031 SSANILLGHTRGATGYFTYHTVLTQTSADFDALKARLIASNT-ETWINGS--AY 1087
QY 1091 EVFPYITNIFYEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISVILV 1150
DB 1088 RVFPYSVFYFYEQYLTIIDTIFNLGVSLSGALFVTLVLLGCELMSAVIMCATIAMLV 1147
QY 1151 DTGVFMALWDISYNAVSLINLVAGVMSVEFVSHITFSFAISKPTWLEBAKATISMG 1210
DB 1148 NMEGVMMWLWGISLNAVSLNVMSCGISVBFCSHITRAFTVSMKGRVERAEALAHMS 1207
```

```
QY 1211 AVFAGVAMTNPGLVGLAKAQIQIFPFRMLNLLTLLGLLHGLVPLPVLVSYGPDVN 1270
DB 1208 SVFSGITLTKFGGIVLAFKASQIFQIFVFRMYLAWVLGATHGLFLPLVLLSYIGPSVN 1267
QY 1271 PA 1272
DB 1268 KA 1269

RESULT 11
Q7TMD4 MOUSE
ID Q7TMD4 MOUSE PRELIMINARY; PRT; 1277 AA.
AC Q7TMD4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE Niemann Pick type Cl.
GN Name=Npci;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Scheetz T.B.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052437; AAH52437.1; -; mRNA.
DR EMBL; BC054539; AAH54539.1; -; mRNA.
DR GO; GO:000687; P.endocytosis; IMP.
DR InterPro; IPR004765; NP_C_type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFBMS; TIGR00917; 2A060601; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1277 AA; 142882 MW; 3B42230AAC8E564E CRC64;

Query Match 34.7%; Score 2400; DB 2; Length 1277;
Best Local Similarity 40.4%; Pred. No. 8.7e-161;
Matches 528; Conservative 233; Mismatches 443; Indels 104; Gaps 27;

QY 11 LWALLRLAQSPFYTIHQPGYCAFYDECGKNPELSSGLMTLSNVSCLSNTPARKITGDH 70
```


QY 577 NNYPADGPELAQKWEAFLEBEMRAFORRMAGMFOVTTAERSLEDEINRTAEDLPF 636
 Db 568 NNYNDTERLQKAHWEKFIPIVFNKYN---PNLTISFTTSTERSIDELNRSNGDIFTV 624
 QY 637 ATSYIVIFYISIALGYSYWSRVMVDKATLGLGVAVVILGAVMAAGFFSYLGRSSL 696
 Db 625 IISYALMFYISIALGHIKSSLLVDSKISGIAGILIVLSVACSLGIFSYVGIPTL 684
 QY 697 VILQVVPFVLSVGADNIFIVLEYQRLPRRGPPEVHIGRALGRVAPSMMLCSLEAI 756
 Db 685 IIVIEVPFLVAVGVNIFILVQYQDERLHGETLDQQLGRVGEVAPSMFLSPSEAV 744
 QY 757 CPELGTALTPAVRFTALTSGLAVILDFLOWSAFVALLSLDSKQESRLDVCCCKPQ 816
 Db 745 AFLGALSOMPVAHTFSLFAGMAVLIDFLQITCFVSLGLDKIKQEKRLDVLCCVRS 804
 QY 817 ELPPPCQ-CEGLLLGPFQKAYAPFLHMTITRGVILLFLALFGVLSYSCHISVGLDQSL 875
 Db 805 EDGTSVQASESCLFLFKHSYSPLLKDMWRIVTAIFGVLSFVAVLNKVEIGLDQSL 864
 QY 876 ALPKDSYLDYFLFLNRYFEVGAIPVYFVTLGYNFSSEAGMNAICSSAGCNNSFTQKIQ 935
 Db 865 SMPDSDYVMDYFKSL-KYLHAGPPVYFVLEEGHDYTSLKQNMVCGMCCNDSLVQQLF 923
 QY 936 YATEPEQSYLAIPASSWDDFDLWLP--SSCRLYISPNKDKPCPTVNSLNLKNCM 994
 Db 924 NAAQLDSYTRIGAFSSWDDYDFMWKPOSSCRVNS---TDRFCNASVDPACIR-CR 979
 QY 995 SITM-GSVRPSVEQPHKLPWFLNDRPNKCPKGLAAVSTSVN-LTSDQGVLASREWAY 1052
 Db 980 PITQEGKQPQGGDFMRFLPMPFLSDNPNPKCGKGGHAAVSSAVNIIIGNDTGVGATYFMTY 1039
 QY 1053 HKPLKNSQDYTEALRAARELAANITADLRKVPCTDPA-PEVFPYTTITNVFYQYLTLPE 1111
 Db 1040 HTVLQTSADFTAMRKANLIASNIT----KTWGLEGSNRYVPYSPYFYFYQYLTIID 1095
 QY 1112 GLFMLSCLVPTFVSCLLGLDLRLSGLNLISIVMLIVDTVGFMAWMDISYNAVSLINL 1171
 Db 1096 TIFNLSVSLGALFLVTIILGCDLMSAVIMCITIAMILVNMFGVWMLGSLNAVSLVNL 1155
 QY 1172 VSAVGSVFEVSHITRSPATSKPTWLERAKATISMGSAVAGVAMTNLPGILVLGLAK 1231
 Db 1156 VNSCGISVFEFCSHITRAFTVSMKGSRAQRAELAHMGSVFSGITLTKFGVIVLAFAK 1215
 QY 1232 AQLIQIFFRLMLITLLGLHGLVPLPVILSVGVDVNPALAL---EOKRAEE 1282
 Db 1216 SQIFQIFYRMYLAWVLLGATHGLFLPLVLLSYIGPSINKAKSLATQEQYKGT 1269
 RESULT 14
 ID NPCL MOUSE
 AC O35604; O35605; STANDARD; PRT; 1278 AA.
 DT 30-MAY-2000. (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-MAY-2005 (Rel. 47, Last annotation update)
 DE Niemann-Pick C1 protein precursor.
 GN Name=Npcl;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=97362324; PubMed=9211850; DOI=10.1126/science.277.5323.232;
 RA Loftus S.K., Morris J.A., Carstee E.D., Gu J.Z., Cummings C.,
 RA Brown A., Ellison J., Ohno K., Rosenfeld M.A., Tagle D.A.,
 RA Pentchev P.G., Pavan W.J.;
 RT "Murine model of Niemann-Pick C disease: mutation in a cholesterol
 RT homeostasis gene.";
 RL Science 277:232-235(1997).
 RN [2]

RP SUBCELLULAR LOCATION.
 RX MEDLINE=99145606; PubMed=9990080; DOI=10.1073/pnas.96.4.1657;
 RA Patel S.C., Suresh S., Kumar U., Hu C.Y., Cooney A.,
 RA Blanchette-Mackie E.J., Neufeld E.B., Patel R.C., Brady R.O.,
 RA Patel Y.C., Pentchev P.G., Ong W.-Y.;
 RT "Localization of Niemann-Pick C1 protein in astrocytes: implications
 for neuronal degeneration in Niemann-Pick type C disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:1657-1662(1999).
 CC -!- FUNCTION: Involved in the intracellular trafficking of
 CC cholesterol. May play a role in vesicular trafficking in glia, a
 CC process that may be crucial for maintaining the structural and
 CC functional integrity of nerve terminals.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Found in late
 CC endosomes and lysosomes.
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in perisynaptic
 CC astrocytic glial processes. Also expressed in heart, spleen, lung,
 CC liver, skeletal muscle, kidney, testis.
 CC -!- INDUCTION: Activated by the drugs progesterone and U-18666A which
 CC block cholesterol transport out of lysosomes and by the
 CC lysosomotropic agent NH4Cl.
 CC -!- DOMAIN: A cysteine-rich N-terminal domain and a C-terminal domain
 CC containing a di-leucine motif necessary for lysosomal targeting
 CC are critical for mobilization of cholesterol from lysosomes.
 CC -!- SIMILARITY: Belongs to the patched family.
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.

 DR EMBL; AF003349; AAB63372.1; -; mRNA.
 DR EMBL; AF003349; AAB63373.1; -; Genomic_DNA.
 DR PIR; T30188; T30188..000024413; Mus musculus.
 DR Ensembl; ENSMUSG000000024413; Mus musculus.
 DR MGI; MGI:1097712; Npcl.
 DR GO; GO:0006897; P: endocytosis; IMP.
 DR InterPro; IPR004765; NP_C type.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; 1.
 DR TIGRPFAM; TIGR00917; 2A060601; 1.
 DR PROSITE; PS0156; SSD; 1.
 DR Glycoprotein; Lysosome; Signal; Transmembrane.
 KW Glycoprotein; Lysosome; Signal; Transmembrane.
 FT SIGNAL 1 23 Potential
 FT CHAIN 24 1278 Niemann-Pick C1 protein.
 FT TRANSMEM 271 291 Potential.
 FT TRANSMEM 352 372 Potential.
 FT TRANSMEM 623 643 Potential.
 FT TRANSMEM 655 675 Potential.
 FT TRANSMEM 685 705 Potential.
 FT TRANSMEM 761 781 Potential.
 FT TRANSMEM 834 854 Potential.
 FT TRANSMEM 1099 1119 Potential.
 FT TRANSMEM 1125 1145 Potential.
 FT TRANSMEM 1196 1216 Potential.
 FT TRANSMEM 1228 1248 Potential.
 FT DOMAIN 621 786 SSD.
 FT MOTIF 1275 1278 Di-leucine motif.
 FT COMEBIAS 250 260 Poly-Pro.
 FT CARBOHYD 71 71 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 123 123 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 138 138 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 186 186 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 223 223 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 229 229 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 415 415 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 460 460 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 479 479 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 525 525 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 474 474 S -> F (in Ref. 1; AAB63373).
 FT CONFLICT 479 479 N -> D (in Ref. 1; AAB63373).

SQL	SEQUENCE	1278 AA; 142889 MW; 43C71CB47D283674 CRC64;
	Query Match	34.5%; Score 2385; DB 1; Length 1278;
	Best Local Similarity	40.2%; Pred.No.1e-159;
	Matches	525; Conservative 233; Mismatches 443; Indels 104; Gaps 27;
QY	14	LLRLAQSEPTTHIQPGYCAFYDECGKNPGLSGLMTLSNVCLSNTPARKITGDHLIL 73
DB	12	LLLLLCPAQVFSQ-----SCVWYGECC---TATGD-----KRYNCKYSGPPKPLPKDGYDL 59
QY	74	LQKICPRLYTGNTQACSAKQOLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSPNQ 133
DB	60	VOELCPGLFF-DNVSLLCCDIQLOLTKNLQLPQLFSRCPSCFYNLMTLFCBLTCSPHQ 118
QY	134	SLFINTRVAQLGAGOLPA----VVAYEAFQHSFAEQSYDSCSRVVRVPAATLAVGTM 189
DB	119	SOFLNVATATEDYFDPKTPENKTNVKELEYVYQGQSPANAMYNACRDVEAPSNKALGLLC 178
QY	190	GVYSGALCNAORWLNFOGDTGNGLAPLDI-----TFHLLBPQAVGSGIQPLNEGVARCN 244
DB	179	GRDARA-CNATNWIEYMNKONGQAPFTIIPVFSDSLIL-----GMEPRNATKGCN 229
QY	245	ESQGDVATCSQDCAASC-----PAJARPOLDSTF-----YLGQMPGSLV 286
DB	230	ESVDEVTGFCQDCSIVCGPKQPQPPPPMPWRIMGLDAMVTVMVTVVAFLVFFFGALL 289
QY	287	LIIILCSFPAVUTILLVGFRAVPARDKMWDPKGTSLDKLSFSTHTLGLQFGQWGT 346
DB	290	AWCHRRRYFVSEYTPIDSNIAFSVNSS-----DKGEASCCDPLGAAPDDCLRRMFTKWGA 345
QY	347	VWASWPLTILVLSVPVVALAAGLVFTLTTPVELMSAPNSQARSEKAFHQHFGPPFR 406
DB	346	FCVRNPTCIIIFSLAFITVCSGLVQVVTNPVELNAPHSQARLEKEYFDKHGPPFR 405
QY	407	TNQVILTAPNRSRYDLSLLGPK-NFSGILDLDLLELLELQERLRHLQWMSPEAQBN- 464
DB	406	TEQLIIQAPNTSVHIYEPVAGADVPPGPPPLNKELTHQVLNLQ-----IAIESITASYN 460
QY	465	--ISLQDICYAPLNPDNTSLVDCINSLLQYFQNNRTLLLTANQTLMGQTSQV----- 516
DB	461	ETVTLQDICVAPLSYNNK---NCTIMSVLNFQNSHAVL-----DSQVGDDPYI 506
QY	517	--DWKDHFLYCANAPLTFKQGTALALSADYDAPVFPFLAIGYKGYDSEAEALIMTF 574
DB	507	YADYTHFLYCYRAASLNDTSLHGPCLTGTFGGVFPVFLVGGYDDQYNNATALVITP 566
QY	575	SLNNYPAGDPLRLAQAKLEEAFLPEMRAPQRMAGMFQVTTFAERSLEDEINRTTAEDLP 634
DB	567	PUNNYNDTERLQRAWAKEPISFVKNYKN---PNLTISFTAERSIEDELNRESNSDVP 623
QY	635	IPATSYIIFLYISIALGYSYSSWRVMSKATILGLGVAVVLGAVMAAGFFSVLGI 694
DB	624	TVIISYVMVFLYISIALGHQICSRLLVDSKISLGIAGILVLSVACSLGIFSYMGNPL 683
QY	695	SLVILQVVPFLVLSGADNIFLVLEYOKRPPRGPPEVHIGRALGRVAPSMILCSLSE 754
DB	684	TLIVIEVIFPLVAVGVNDIFLVQTYQDRLEQETLDQGLRILGEVAPTMFLSSFE 743
QY	755	AICFFLGLATPMPAVRTFALTSLGLAVILDFLLQMSAFVALLSDSKRQEAERLDVCCCVK 814
DB	744	TSAFFFGALSSMPAVHTFSLPAGMAVLIDFLLIQITCFVSLGLDIDKROEKHLDILCCVR 803
QY	815	POELPPPGG-----EGLLGPFPQKAYAPFLHWHITRGVULLFLALFGVSLYSMCHISVG 870
DB	804	GAD-----DQGGSHASESYLFRFPKNTFAPLLKDWLRPIVAVFVGLVFSVAVNKNVDIG 860
QY	871	LQELALPKDYLVDYFLNRYFEGAPVYFVTLTGNFSEAGNAICSSAGCNFSF 930
DB	861	LQSLSMFNDSVIANFNSLAQYLIHSPPVYFVLREGYNSRKNQNVCGMGCDNDSL 920
QY	931	TQIKQYATEFPQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKDKFCPSTVNSLNC 989
DB	921	VQOIFNAELDTYTRVGFPAPSSWIDDYFDWVSPQSCCRLY---NVTHQFCNASVMDPTC 977

QY	990	LKNCMSIT-MGSVRPSVEQFHKYLWFLNDRPNIKCPKGGLAAYSTSVNLTSQGVLA-AS 1047
DB	978	VR-CRPTPEGKQRPQGEKFKFLPMFLSDNPNPKCGKGHAAYGSAVNIVGDDTYIGAT 1036
QY	1048	RFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAFEVFPVYTIINVFYEQILT 1107
DB	1037	YFMTYHTILKTSADYTDAMKKARLIASNITETMRS-KGSD--YRVFPYSVFVFPVEQYLT 1093
QY	1108	ILPGLFWLSLCLVPTFAVSCLLIGLDI-RSGLNLLSIVMLVDTVGFMAWLDISYNAVS 1167
DB	1094	IIDDTIFNLVSLSGIFLVTLWLGCLWSAVINCITAMILVNMFGVMWMLWGLSNAVS 1153
QY	1168	LINLVSAVGSVFVSHITRSFALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVL 1227
DB	1154	LVNLMVCGISVEFCSHITRAFTWSTKGSVRSAEALAHMGSSVFSISITLTKFGIIVL 1213
QY	1228	GLAKAQIQLIFFPRLNLLITLGLLHGLVFLPVLVSYYVGPVNPA 1272
DB	1214	AFASQIFEIFYFPMYLAWLGLGATHGLIFLPLVLSYIGPSVKA 1258
RESULT 15		
Q9GLC9 BOVIN		
ID	Q9GLC9 BOVIN PRELIMINARY;	PRT; 1277 AA.
AC	O9GLC9;	
DT	01-MAR-2001 (TrEMBLrel. 16, Created)	
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)	
DE	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Niemann-Pick type CI disease protein.	
GN	Name=NPCI;	
OS	Bos taurus (Bovine).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;	
OC	Pecora; Bovidae; Bovinae; Bos.	
OX	NCBI_TaxID=9913;	
RN	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RC	TISSUE=Aorta endothelium;	
RA	Murakawa M., Freeman M.W.;	
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AF299073; AAG24620.1; -, mRNA.	
DR	GO; GO:0016021; C:integral to membrane; IEA.	
DR	GO; GO:0005764; C:lysosome; IEA.	
DR	GO; GO:0016020; C:membrane; IEA.	
DR	GO; GO:0008158; F:hedgehog receptor activity; IEA.	
DR	GO; GO:0030301; P:cholesterol transport; IEA.	
DR	InterPro; IPR004765; NP_C.type.	
DR	InterPro; IPR003392; Patched.	
DR	InterPro; IPR000731; SSD 5TM.	
DR	Pfam; PF02460; Patched; I.	
DR	TIGRFAMs; TIGR00917; 2A060601; 1.	
DR	PROSITE; PS0156; SSD; 1.	
SQL	SEQUENCE 1277 AA; 141911 MW; 30C91ABFD1ABF1CF CRC64;	
Query Match		
Best Local Similarity 34.3%; Score 2367.5; DB 2; Length 1277;		
Matches 517; Conservative 227; Mismatches 445; Indels 101; Gaps 26;		
QY	33	CAPYDECKNPGLSGLMTLSNVCLSNTPARKITGDHLILQKICPRLYTGNTQACCS 92
DB	25	CIMWGECC---IASGD---KRYNCRYSGPPEPLPDQGYDLVQELCPGFFFG-NVSLCCD 76
QY	93	AKQLVSLASISITKALLTRCPACSDNFVNLHCHNTCSPNOSLFINTRVAQL---GAGQ 149
DB	77	VQQLTKDNLQLPQLFSRCPSCFYNLVNLFCBLTCSPROQFLNVATATEDYVDATNQ 136
QY	150	LPAVVAYEAFY-QHSFAEQSYDSCSRVVRVPAATLAVGTMCGVYSALCNAORWLNFOGD 208
DB	137	TKTNVKELOYVGSFANAMYNCSRDVEAPSNKALGLLCREASA-CNATNWIEYMFN 195
QY	209	TGNGLAPLDIT--FHLLEPGQAVGSGIQPLNEGVARCNESQGDVATCSQDCAASCPAI 266

Db 196 KONGQAPFTTIPVFSDLPT-----HGMEMPNNATKCDSESDVDEVTGCPSCQDCSACVCGPK 250
QY 267 ARPQ-----ALDSTYLGOMPGSLVLIILCSFVAVVTLILVG--FRVAPARDK-- 313
Db 251 PQPPPPPPVWKILGLDAMY-----VIMSTYMAFLLVFFGAFFAVCMVKRYF 298
QY 314 -----SKMVDPKKGTSLSKLSFSTHTILGQFFQCGWGTWVASWPLTILVLSV 360
Db 299 VSEFTPIDGNI PFSINADKGGTCCDPLGAAPAEHLRLFLFEMWGSFCVRHFGCVVFFSV 358
QY 361 IPVVALAAGLVFELTTPDVELMSAPNSQARSEKAFHQHFGPFFRTNQTIVILTPNRSSY 420
Db 359 AFIAACSSGLVFIQVTTDPDLMSAPGSAQLEKEYDFTHFGPFFTEQLIIRAPTPPH 418
QY 421 RYDLSLLGPK-NFSGILDLDDLELELERLHLQVMSPEAQRNLSQDICYAPLNPN 479
Db 419 IYEPYPSGADVPEGPLAVNHLVLDLQTAIE--SITASYNNTVTLRDIICVAPLSPN 476
QY 480 TSLYDCCINSLLQYFONNRTLLLTANQTLMGOTSOV--DWKDHFLYCANAPLTKDGTG 537
Db 477 ---QNCITLSVLNYFQNSHVL-----DHQVGDDFFVYADYTHFLYWRAPASLMDTSL 528
QY 538 LALSCHADYGAPEPPLAIGGVKDYSEAEALIMTFSLNYPAGDPRLAQAQKLEEAFL 597
Db 529 LHDPCLGTEGGVPFVPLVGGDDQNNATALLVTFPVNNYINDEKLRQAQWKEKFI 588
QY 598 EEMRAFQRMAGMFOVTFATERSLEINRTAEDLPIFATSYVIFLYISIALAGSYSSW 657
Db 589 NFQVQYEN---PNLITSFKAERSIEDENRESNDVFTVLISYGVWFLYISIALGHKSC 645
QY 658 SRMVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSVLIVQVFPFLVLSGADNIFIF 717
Db 646 RELLVDSKILLGAGVLI VLSPVACSLGFSYWGSPPLTLIVIEVIFPLVAVGVNDIFIL 705
QY 718 VLEYQRLPRPGEPRVHIGRALGRVAPSMILCSISEALCFELGALTMPAYRTALTSG 777
Db 706 VQTYQIDELQOGETLDDQGRVLGEVAPSMFLSSFAETVAFPLGGLSVMPVHTLSLFA 765
QY 778 LAVILDFFLQMSAFVALLSLSKQREASRLDVCCKVKQBELPPPG-----QOEGILL 828
Db 766 MAVLIDFLQITCFVSLGLDVKRQEKQDLVLCV-----GGAADDAGIQASESL 817
QY 829 LGFFOKAYAPFLHNTTRGVLLFLALFGVSLYSNMCHISVGLDQBLALPKDSYLLDYL 888
Db 818 PRFRNSYAPLLKDWMRPLVAVFVGLSFAVLNKEVIGLDQSLSPDSDSYVTDYQ 877
QY 889 FLNRYFEVGAPEVYVFTLGYNFSSEAGMAICSSAGCANNFSTQKIQYATEPPEQSILAI 948
Db 878 SLNQYLHAGPPYVFLVEEGHDTSTKQNMVCGGLGCNNDSLQVQVFTAAQLDSYTRIGF 937
QY 949 PASSWVDDFDLWLP--SSCRLYISGPNKDKFCPTSVNSLNCNKMCSIT-MGSVRPSVE 1006
Db 938 APSSWIDDYFDWVKPOSSCCRIYNS---TEQFCNASVNVPTCV-CRPLTPGKQRPQGA 993
QY 1007 QPHKYLPLWFLNDRPNIKCPKGLAAYSTSVNLTSDG-QVLASRFMAVHKPLKNSODYTRA 1065
Db 994 DFMRFLPMLPSDNPNPKCKGKGHAAVSAAVNILDNGTRVGATVFWYTHVLTQTSADFIDA 1053
QY 1066 LRAARELANIITADLAKVGTDPAEVFPYTTNTVFEYQYLILPGLPMLSLCLVPTFA 1125
Db 1054 MEKARLIASNTITWNNQOQGD---HRVFPYSVFPYFEQYLTMIDDTIFNLVSLGAIFL 1110
QY 1126 VSLCLGLDLRGLNLLLSI VMLVDTVGFMAIWDISYNAVSLINLVSAGVMSVRFVSHI 1185
Db 1111 VAVVLGGLSNVAVINCATIAMLVNMFGVMMGLSNVSLVNLWMSGSLSVFSCSHI 1170
QY 1186 TRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLIVLGLKAQILQIIFFRNLNL 1245
Db 1171 TRAFVTSTKGSVERAEALSHMGSSVFSGITLTKFGGIIVLAPAKSQLFIQFYFRWYLA 1230
QY 1246 IYLLGLLGLVFLPVLSTVSGDDVNDALAL 1275
Db 1231 MYLLGATHGLIIFLPVLLSYIGPSINKAKSL 1260

RESULT 16

Q9TT75 RABIT PRELIMINARY; PRT; 1286 AA.
AC Q9TT75_ RABIT PRELIMINARY; PRT; 1286 AA.
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Niemann-Pick type C1 disease protein.
GN Name=NPC1;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Murakawa M., Freeman M.W.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF202730; AAF18444.1; -, mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:heghehog receptor activity; IEA.
DR GO; GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1286 AA; 142723 MW; 9266CC310F64CCD7 CRC64;

Query Match 34.2%; Score 2361; DB 2; Length 1286;
Best Local Similarity 39.0%; Pred. No. 5.1e-158;
Matches 513; Conservative 237; Mismatches 455; Indels 110; Gaps 26;

QY 7 RGMILLALLRLAQSEPYTHIQPGVCYAFYDSCGNKPELSGSLMTLSNVSCLSNTPARKI 66
Db 13 RGLLGLLGLLGLLCPAQL-----PQSCWYGEQ---IASGD---KRYNCQYSGPKPL 60
QY 67 TGDHILILQIKCPRIYTGNTQACCSAKQLVSLASLSITKALLTRCPACSNFVNLHCH 126
Db 61 PKDGYDLIQELCPGFF- DNVSLCCDVQQLQTLKDNLQLPLQFLSRCPSCFYNLMLFCE 119
QY 127 NTCSPNQSILFNVTRVAQLGACQLPAVAY-----EAFYQHFABQSYDYS 171
Db 120 LTCSPRQSLNVT-----ATEAYVDPATNQTNTYKELQYIYIGORPADAMYNA 168
QY 172 CSRVVRPAAATLAVGTMCGVYGSALCNQORWLFQDGTGNGLAPLDIT--FHLLEBPQAV 229
Db 169 CRDVEAPSSNDKALGLLCKGDANA-CNATNWTYEMFDRKNGQAPFTITPISDLPV---- 223
QY 230 GSGIOPLENGVARCNESQGDVATCSQDCQCAASCPAIARPQ-----ALDSTFYLQ 280
Db 224 -RGMEPMNATKGCNEAVDEVTGPCQDCQCSVVCQPKQPPPPPIPRIFGLDAMYVIMW 282
QY 281 MPGSLVLIILCSVPAVTVILLVGRVAPARDKSKMVPDKGTSLSDKLSFSTHT----- 335
Db 283 ITYMAFLPMFFGTFPAVMCY-----RKRYFVSEYTPIDSNIAFVSNTDKGE 329
QY 336 -----LLGQFFQG-----WGTWVASWPLTILVSLVVALAAGLVFTSLTTPDVEL 382
Db 330 ASCCDLLGAAPGCLURRLFTRMGSPCVRNPGCVIFFSGLFIACSSGLVYVVTNPIDL 389
QY 383 WSAPNSQARSEKAFHQHFGPFFRTNQVILTPANRSSRYDLSLLGPK-NFSGILDLDDL 441
Db 390 WSAPSSQARQDKKFFDAHFGPFFRTEQLIIRAPHTNKHTYQYPYPSGADVFPGLDKQL 449
QY 442 LELLEQLERLHLQVMSPEAQRNLSQDICYAPLNPNDTSLVDCCINSLLQYFQNNRTLL 501
Db 450 HQVLDLQTAIE--DITASYNNTVTLQDLCIAPLSFYNK---NCTILSVLNTFQNSHML 504

```
QY 502 LLTANTQILMGTSQVDMKDFLYCANAPLTPKDGATLALSCMADYGAPVPPFLAIGYKG 561
Db 505 DHEQDDFF---VYADHITHELVCVRVPASLNDTSLLDHPCGLTGGFGGVPFWLVGLGYDD 561
QY 562 KYSEAEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVTTAERSL 621
Db 562 QYNNATATLVITFPVSNYNDTEKLQAKAWEKFINPVKNYEN---PNLTISTAERSI 618
QY 622 EDEINRTTAEDLPATFATYIVIFLYISALGSYSWSRVMVDSKATLGLGVAIVVLGAVM 681
Db 619 EDELDRENSDVFTVVISYAVWFLYISALGHKSCRFVDSKSLGIAGILVLSVA 678
QY 682 AAMGFYSYLGRSSILVQVPPVLVSVDGANIFVLEYQLRPRRPGEPREVIHGRALG 741
Db 679 CSLGIFSYIGIPLTLIVIEVIFPLVAVGVNIFILVQTYQRDERLOGETLDQGLRVLG 738
QY 742 RVAPSMLLCSLSEATCFPLGALTMPAVRTALTSLGLAVILDFLQMSAFVALLSLDSKR 801
Db 739 EVAPISFLSSPAETVAFVFLGLSVNPAVHTTSLFAGMAVFDLLOITCFVSLGLDIDR 798
QY 802 QASRLDVCCCKVQELPPGQ-GEGLLGGFKQAYAPPLHWTIRGVVLLLFALFQVS 860
Db 799 QEKQLDLVLCPRGAEAGTSIQASSESYLFRFPKNSYSPLLLKDMWRPIVIAVFGVLSFS 858
QY 861 LYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPVYVVTTLGYNFSEAGMNAIC 920
Db 859 IAVLNKVEIGLQSLMPDSDSYVDYFKSLGQYLHAGPPVYVFLBEGHNYTSLQGNMVC 918
QY 921 SSAGCNNEFTOKIQVATFEFQSYLATPASSWDDFDLWLT-SSCRLYISGNKDKF 979
Db 919 GGLGCDNLSVQOIFNAAQLDNTYRIGAPSSWIDDYFDWVQPSCCRV---SNVTQF 975
QY 980 CPSTVNSLNLCKNCSIT-MGSVRPSEVQFHKYLPWFLNDRPNIKCPKGGAAYSTSVNL 1038
Db 976 CNASVVDPAVCR-CRPLTPEGKORQGEDFMFLPWLSDNPNKCGKGGHAAVGTAVNI 1034
QY 1039 TSDG-OVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVPPYTI 1097
Db 1035 LGNGTEVGATYFTMTHTVLOTSAFDIDAMKKAQLIAGNATM-GLKGS--RCRVFPYSV 1091
QY 1098 TNVFEQYITLPEGLFMLSCLVPTFAVSCLLGLDLASGLNLISIVMLIVDTVGEMA 1157
Db 1092 FYVFEQYITLDDAVFNGLSVLGAVFVAVVAVVLLGCELWSAVIMCVTITMILVNMFGVM 1151
QY 1158 LMDISYNAVSLNLSVAGSVSEFVSHITRSPASTKPTWLERAKEATISMGSAVPAGA 1217
Db 1152 LWGISLNAVSLNVLNMSCGISVEFCSHITRAFTVSGKSRVARAEALAHMGSSVFSGIT 1211
QY 1218 MTNLPGLVILGLAKAQIIFEPRLNLITLGLLGLVFLPVILSYVGPVNP 1272
Db 1212 LTKFEGIVVLAFAKSQIFQIFPRMYLAMVLLGATHGLIFLPLVLLSYIGPSTNKA 1266
```

RESULT 17

```
Q4RW5_TETNG PRELIMINARY; PRT; 1209 AA.
ID Q4RW5_TETNG PRELIMINARY;
AC Q4RW5;
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14981, whole genome shotgun sequence.
DE (Fragment).
DE ORFNames=GSTENG0027641001;
OS Tetraodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN NUCLEOTIDE SEQUENCE.
RP Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA
```

```
RA Mauceli B., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dosat C., Segurens B.,
RA Basilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellani V., Katinka M., Vacherie B.,
RA Blomont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Creaud C., Lardier G., Brottier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chappie C., McKernan K.J., McEwan P., Bosak S.,
RA Kallie M., Volff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolious H.;
RT "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014981; CAG07097.1; -; Genomic_DNA.
FT NON TER 1
FT NON TER 1209
FT SEQUENCE 1209 AA; 133913 MW; 2046F9D3DE4E48F1 CRC64;
```

Query Match 33.8%; Score 2338; DB 2; Length 1209;
Best Local Similarity 41.5%; Pred. No. 2e-156;
Matches 520; Conservative 217; Mismatches 414; Indels 102; Gaps 30;

```
QY 72 ILKQICPRLYTGPTQACCSAKQLVSLASLTKALLTRCPACSNFVNLHCHTCTSP 131
Db 1 VCLQELCPG-YDYNRSLLCCDQNLHTLKEVLPQLSRCPACFFNLNVLNLCETCTSP 59
```

```
QY 132 NQSLFINTRVAQAGOLPAVAYEAFYQHSFAEQSDSCSRVRAAATLAVGTWGV 191
Db 60 HQSQFMATKLSG-----PDWEVQYIIGLTFANMTNACRDVQAFSSNVKALSLLCG- 112
```

```
QY 192 YGSALCAORNLNFGDGTGNGLAPLDTFHLLEP--QOAVGSGIQPLNEGVARCNESGD 249
Db 113 KDAKCHATNWIQWNTDNEQAPFIT-----PIFSDVPVSGVTPMNDTAACTDGLD 167
```

```
QY 250 DVATCSCQDCAASC-PATARPQALDSTFYLGQMGSLVLIILCSVPAVTVILLVGRVA 308
Db 168 GSGPCSCQDCTNACGPRVPPPTPAWKILGM---DAMTVIMWFSYMAFLIFVGSLLIA 224
```

```
QY 309 PARDK-----SKMVDPKKTSLS---DKLSF-----STHTLLGQFFQG----- 343
Db 225 WCHRKETTMSYGPFLDSKNRPSLNRPDPVHPFILLDDASCCTELGERFESYLRSCFSC 284
```

```
QY 344 WGTWVASWPLTILVLSVIPVVALAAGLVFTBLTDPVELWSAPNSQARSEKAFHQHFGP 403
Db 285 WGSFCVLPVCLVGLSLVAVSSGGLIYRITTDVPLWSPSSQARQERYFDSHFPG 344
```

```
QY 404 PFRNQVILTAPNRSYRYDSLLGPK-NFSGILDLLELLELLELLELLELLELLELLEL 462
Db 345 PFRTAQLIITSPNDTFLYSPVMGPDIPFKAVLSKDLHQVLDLQLDIESL-VAYEQG 403
```

```
QY 463 RNISLDICAPLNPNDTSLVDCCLNSLQYFQNNRTLLLTANOTLMGQ-TSQVDWFKDH 521
Db 404 -SVTLKDICLAPLSYND---NCTLSLVNLFQNSHS-----TLDHVLKDEFLVYADFHS 455
```

```
QY 522 FLYCANAPLTPKDGATLALSCMADYGAPVPPFLAIGYKGKDYSEAEALIMTFLSL-NNYP 580
Db 456 FLYCVSAPASLNDTTLPHDPCLTGFGVPFWPALGAGYDDTNNYNNATLAVVTFPINNNY- 514
```

```
QY 581 AGDP-RLAQAKLWEEAFLEEMRAFORRMAGMFQVTTAERSLEDEINRTTARDLPIFATS 639
Db 515 --DPTKGLKTLAWEKEFRFMKNYSN---PNUTIAFSAERSIEDEINRESNDISTIVVS 569
```

```
QY 640 YIVIFLYISLALGSYSWSR-----VMVDSKATLGLGVAIVVLGAVMAAMGFPSY 689
Db 570 YIMFVYISLALGHITQSFRLPLPHVLLLLLLLVDSKVSIGILVLSVSSSLGIFSY 629
```


DR FlyBase; FBgn0024320; NPC1.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0005764; C:lysosome; IEA.
DR GO:0016020; C:membrane; IEA.
DR GO:0008159; F:hedgehog receptor activity; IEA.
DR GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1287 AA; 143290 MW; 540AAB292A37790C CRC64;

Query Match 32.7%; Score 2262; DB 2; Length 1287;
Best Local Similarity 37.8%; Pred. No. 5.3e-151;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LLWALLRLAQSEPTTIHQPGYCAFYDECGNPELSGLMTLSNVSCLSNTPARKITGD 69
DB 17 LLAALFTLIQSKD-----CVYGVCTN-DPSHS-----QNCYPNGTAKEMATD 62
QY 70 HLILQKICPRLYTGENTQACSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNWC 129
DB 63 GLELLKRCGFLSENKFCCKNQVELLNKNVLAGNILDRCPCSMENLVRIHCQPTC 122
QY 130 SPNQSFLNTRVAQAGQLPAVVAEYAFYQHSFAEQSYDSCSRVRPAAATLAVGTMC 189
DB 123 SPQAEFMHVATQKNKKGD-EYISSVDLHISTEYINKTYKSCQSVSPQGLAFDLMC 181
QY 190 GYVGSALCNAQRLNFOGDTGNGLAPLDTIPLLEPGQAVSGGIQPLNEGVARNCSOGD 249
DB 182 GAYSARCNPTKFNFMGDATAHYVPFQIITYQHEP-KSNSNFTPLNVTTPCNAQVSS 240
QY 250 DVATSCDCDCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSVPAV-VT 299
DB 241 KLPACSCDCLSCPGPPEPRPPEPKIVGLDAVF-----VMAAFLVGLV 288
QY 300 ILLVG---FRVAPARDKSNVDPKGTSLSKLSFS-----THLLGQFFQGW 344
DB 289 VFLMGSLFTQGSMDNDNFQVD---GNDVSDMPYSENDSYFEKLGATHTETLETFYTKW 345
QY 345 GTWVASMPLTILVSVIPVLAAGLVFTELTDPVLEWSPNSQARSEKAFHDQFGPF 404
DB 346 GTYFASNPGTLIAGASLVILGYGINFIEITDPVKLWASPNKSRLEREFDFTKFSF 405
QY 405 FRTNQVILTAPNRSSVRYDLSLLGPKNFSGIIDLLELLELQERLRLHQLWSPQAQRN 464
DB 406 YRLQIIIAKAVNLPQIVHNT-SNGPYTFGVPFVDFREPLTKVLDLQEGIKEINA-----NG 458
QY 465 ISLQDICYAPLNPDTSL--YDCINSLLQYQNNRTLLLTANQTLMGQTSQVDWKDH 522
DB 459 TQLKDICYAPLSDGSEIDVSCVQSIWGYGDDRERL-----DDHDENGFNVTYLDAL 514
QY 523 LYCANAPLTPKDGATALASCMDYAGVPFPFLAIGY-----KGKDYSEAEALIMTF 574
DB 515 YDCISNPLY-----CLAYGPGVPDAIALGGLFPDQGLTCKTFELANAILTF 564
QY 575 SLNNYPAGDPRLAQAKLWEAFLEEMRAP-QRRMAGMFQVTFPAERSLDEINRTAEDL 633
DB 565 LVKNH-HNKTDLNALTWEKKFVEFMTVNTYKNNMSQYMDIAFTSERSIEDLNRESQSDV 623
QY 634 PIFATSVIVFLYISLALSGYSWSRWVDSKATLGLGVAVVLGAVMAAMGPFSLGHR 693
DB 624 LTLVSLVIMFMXIALSHGVKFKRVFIDSKITLGLIGVVIIVLASVSVSGVFGYIGLP 683
QY 694 SSLVILQVFPVLVSGADNIFIVLEYQRLPRRPGEPREVHIGRALGRVAPSMLLCSLS 753
DB 684 ATLIIIVEIIFVLVAVGVNDIFILVQTHQDRQKPNETLEQQVGRILGKVPKSMLLTSL 743
QY 754 EATCFEFLGALTPMPAVTTPALTGSLVILDFILQMSAFVALLSILDSKQEAESLQVCCCV 813
DB 744 ESFCFFLGLSDMPAVRAFALYAGVALIIDFLIQITCFVSLFTLTKRREENRMDICCFI 803

QY 814 KPOELPPPGQEGLLGLGFFQKAYAPFLHLHWITRGVVLLFLALFGVSLVSMCHISVGLDQ 873
DB 804 KGKKPDSITISNEGLLYKFTSSVTVFPMKKIVRASVWVIFPFWMLCFSLAIPRIDIGLDQ 863
QY 874 ELALPKDSYLLDVLFLNRYFEFGVAPYFVTTLYGYNFSSSEAGNAVATCSSAGCNFSFTQK 933
DB 864 ELAMPQDSFVLHYFQSLNENLNIGPPVYFVLKGLDAYTNSDQNLVCAGQYCNDDSVLTQ 923
QY 934 IQYATEPPEQSYLAIPASSWVDDFDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNL 990
DB 924 IYLAHSNNTYIARPASSWIDDIYFDMAAASCCCKY-----RKDSGDFCPH--QDTSL 976
QY 991 KNCMSITMGS-VRPSVEQFHKLYFWFLNDRPNIKCPKGGGLAAYSTSVNLTSDDQVL---A 1046
DB 977 R--CNITKNSLLARPEKEFKVLYLPFFLKNDPDDTCAKAAHAYGAVRYSNSHERLNIEA 1034
QY 1047 SRFMAYHKPKNSQDYTEALRAARELANATADLR-----KVP-GTDPAPFEPYPTITN 1099
DB 1035 SYFMAYHTILKSSADYFLALESARKISANTQMLQGRLMNGVPMASALTVEVPYSVY 1094
QY 1100 VFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLLNLISIVMLVDTVGFMAW 1159
DB 1095 VFYEQYLTWMSDITLQSMGIVSLISFVTVFVLMGFVHSALVVVITITWIVNVLGGLMYW 1154
QY 1160 DISYNAVSLINLSAVGMSVEFVSHITRSPATSTKPTWLERAKEATISMGSAVPAVAMT 1219
DB 1155 NISLNAVSLVNLVAVGISVEFCSHLVHSPATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214
QY 1220 NLPQILVGLAKAQLIQIFFRLNLLITLGLLGLVFLPVILSVYVGPVDPVNPALAEQKR 1279
DB 1215 KFAGILVAFASQIFQVYFRMYLGIWIGAAGHLIFLPVLLSYIIGAPVSNARLYHSQ 1274
QY 1280 A---EEAVALVM 1288
DB 1275 AAETHALAGIL 1287

RESULT 19

QYU5W1 DROME PRELIMINARY; PRT; 1287 AA.
AC QYU5W1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NPC1 protein.
GN Name=NPC1; ORFNames=CG5722;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zhang Y.Q., Xue H., Broadie K.S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ249606; CAB56505.1; -; mRNA.
DR FlyBase; FBgn0024320; CG5722.
DR FlyBase; FBgn0024320; NPC1.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0005764; C:lysosome; IEA.
DR GO:0016020; C:membrane; IEA.
DR GO:0008159; F:hedgehog receptor activity; IEA.
DR GO:0030301; P:cholesterol transport; IEA.
DR InterPro; IPR004765; NP C type.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1287 AA; 143277 MW; 8E154FE5E5DEF556 CRC64;

Query Match

32.7%; Score 2256; DB 2; Length 1287;

Db	63	GLELLKRCGFLLENSSENKFCDDKXQVELLNKNVELAGNILDRCPSKMWELVRHICQFTC	122
Qy	130	SPNOSLFNTVTRVAQLGAGQLPAAVAYEAFYQHSFAEQSYDSYSCSRVRVPAAATLAVGTCM	189
Db	123	SPKQAEFWHVATQKXKGD-EYISSVDLHSTIEYINKYKSCQSVQPTQGLAPFLMC	181
Qy	190	GUYGASALCNAQRWLNFGDGTONGIAPLDITPHLLPQQAQVSGIGIPLNEGVARCNESQGD	249
Db	182	GAYSASRCNPTKWFNFMGDATNPVVPQITVIQHEP-KSNSNNFTPLNVTTPVCNQAVSS	240
Qy	250	DVATCSCODCAASC-----PATARQ-----ALDSTEYLGQMPGSLVLIILILCSFEAV-VT	299
Db	241	KLPACSCSDCLSPQGPPEPRPEPKIVGLDAYF-----VMAAVFLVGVL	288
Qy	300	ILLVGV---FRVAPARDKSKVDPKKGTSLSKLSFS-----THLLGQFPQGW	344
Db	289	VFLMGSLFTQSSWDDNFQVD---GNDVSDMPYSENDSYFEKLGATHTFTLFTFTKW	345
Qy	345	GTWVASWPLTILVLSVIPVVALAAGLVTELTTPDVELWASAPNSQARSEKAFHQHGGPP	404
Db	346	GTYFASNPGLTILAGASLWILGYGINFIETITDPVKLWASPNKSLREFFFTKSPF	405
Qy	405	FRTNQVILTAPNRSSRYRDSLLGPKNFSGILDLDLLELELOERLRLHQLQWSPQAQRN	464
Db	406	YRLEQIIKAVNLPQIVHNT-SNGPYTFGPVDFREFLTKVLDLQEGEKEINA-----NG	458
Qy	465	ISLOQICYAPLNPDMTSL--YDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDH	522
Db	459	TOLKIDICVAPLSDGSEIDVSCVQSIWGYGDDRELR-----DDHDEBDGNFVNYLDAL	514
Qy	523	LYCANAPITFDGDTALASCMADYGAQVPFPLAIGY-----KGXDYSEAEALIMTF	574
Db	515	YDCISNPYL-----CLAPYGGPVDPAIALGGLFPDGDQLTGSTKFFELANAIIITF	564
Qy	575	SLNNYPAGDPRLAAQKLEAEFLEMBRAF-QRRMAGMFQVTFPAAERSLEDEINRTAEDL	633
Db	565	LYKNH-HNKTDLLENALTWEKKEVEFWNTYTKNNMQYMDIAPTSERSIEDELINRESQSDV	623
Qy	634	PIFATSYIVIFLYISIALGYSYSSWRVMDVSKATLGLGAVVLGAVMAAMGFFSYLGIR	693
Db	624	LTLVSYLIMFWYIAISLGHVKEKRVFIDSKITLIGIGVILVSLAVSSVGVFGYIGLP	683
Qy	694	SSLVLQVVPFLVLSVGADNIIFVLEYQRLPRRGPCEPREVHIGRALGRVAPSMLLCSLS	753
Db	684	ATLIIIVEIPFLVAVGVNDIFILVQTHQDQKNETLEQQVGRILKGKVGPSILLTSL	743
Qy	754	EALCPFLGALTMPAVRTEFALTSLGLAVILDFLQMSAFVALLSLDSKQEAASRLDVCCCV	813
Db	744	ESFCFLGLSDMPAVRAPALVAGVALIIDFLQITCFVSLTLDTKREENRMDICCFI	803
Qy	814	KPQELPPPGGELLGLGFQKAYAPFLHWTIRGVVLLFLFLPGLVGLYSMCHISVGLDQ	873
Db	804	KGKKPDSITSNELLYKPFSSVYVPLMKKIVRASVMVFFAWLFCFSAIAPRIDGLDQ	863
Qy	874	ELALPKDSVLLDPLFLNRYFVGAVYFVITLGYNFSSEAGNNAICSSAGCNNSFTQK	933
Db	864	ELAMPQDSFVLHYFQSLDENLNIQPPVYFLKGDLAYTNSDQNLVCAQYCNDSVLTK	923
Qy	934	IQVATEFPQSYLAIPASSWDDFDIDW-LTPSSCCRLYISGPNKDK--PCPSTVNSLNC	990
Db	924	IYLASHRSNQVYIARPASSWIDDYFDWAAAASCCYK-----RKDSGDFCPH--QDTSCL	976
Qy	991	KNCMSITMGS-VRPSVEQFHYLPWFINDRPNTKCPKGLLAYSTSVNITSDQVIL---A	1046
Db	977	R--CNITKSNLLRPEEKFEVKYLPFLKONPDOTCAKAGHAAYGGAARYYNSHERLNIEA	1034
Qy	1047	SRPMAYHKPLKNSQDYTEALRABELAANITADLR-----KVP-CTDPAFEVFPYITN	1099
Db	1035	SYPMAYHTILKSADYFLAELARKISANITQMLQGRLMNGVPMASALTVEVFPYSVY	1094
Qy	1100	VFYEQYLITPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLSITVMILVDVTGPFMAW	1159
Db	1095	VFYEQYLITWMSDTLQSMGISVLSISFVVTVFLMGFDVHSALVVVVITITIMIVNVLGGLMYW	1154
Qy	1160	DISYNAVSLINLVSAYGMSVEFVSHITRSPATSKTWTLEAKEATISMGSAVFAGVAMT	1219
Db	1155	NISLNAVSLNVLNVMAGVISVEFCSHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT	1214
Qy	1220	NLPGLIIVLGLAKAQLTIQIFFRNLMLITLLGLHLGLVFLPVLISYVGPDPYNPALALEQKR	1279
Db	1215	KFAGILVLAFAKSIQIQVFFRYMYLGIIVVIGAAKGLIFLFPVLLSYIGAPVSNARLRYHSQ	1274
Qy	1280	A-----EEAVAAVM	1288
Db	1275	AAAEHETALAGIL	1287
RESULT 21			
Q7Q409 ANOGA			
ID	Q7Q409	ANOGA PRELIMINARY;	PRT; 1291 AA.
AC	Q7Q409;		
DT	01-MAR-2004	(T-EMBLrel. 26, Created)	
DT	01-MAR-2004	(T-EMBLrel. 26, Last sequence update)	
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)	
DE	ENSANGP0000010415	(Fragment).	
GN	ORFNames=ENSANGG0000007926;		
OS	Anopheles gambiae str. PEST.		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Culicidae;		
OC	Anophelinae; Anopheles.		
OX	NCBI_TaxID=180454;		
RN	NCBI	TaxID=180454;	
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=PEST;		
RG	The Anopheles gambiae Sequence Committee;		
RT	"Anopheles gambiae re-annotation."		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
[2]			
RP	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=PEST;		
RG	The Anopheles gambiae Sequence Committee;		
RL	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.		
-!- CAUTION: The sequence shown here is derived from an			
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is		
CC	preliminary data.		
CC	EMBL; AA0801008964; EAA12360.2; -; Genomic DNA.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0005764; C:lysosome; IEA.		
DR	GO; GO:0016020; C:membrane; IEA.		
DR	GO; GO:0008158; F:hedgehog receptor activity; IEA.		
DR	GO; GO:0030301; P:cholesterol transport; IEA.		
DR	InterPro; IPR004765; NP C type.		
DR	InterPro; IPR003392; Patched.		
DR	InterPro; IPR000731; SSD 5TM.		
DR	Pfam; PF02460; Patched; 1		
DR	TIGRFAMs; TIGR00917; 2A060601; 1.		
DR	PROSITE; PS50156; SSD; 1.		
FT	NON TER		
FT	1		
SQ	SEQUENCE	1291 AA; 142230 MW; 244A70BEDDA3F519 CRC64;	
Query Match			
Best Local Similarity			
Matches 491; Conservative 242; Mismatches 470; Indels 110; Gaps 31;			
Qy	33	CAYFDECCGNPELSGLMTLSNVSLNTPARKITGDHLLILQKICPRLY-----TGPNTQ	88
Db	8	CWVYIC--NVDSLGRAQ-----YCPYNGTAKAIDTEASDLLKWKCKLLVEDATG-SIN	59
Qy	89	ACCSAKQLVSLSEASITKALLITRCPACSDNPNVNLHCHTCSNQSFLFNITRV-----	142
Db	60	TCCDAQVKVLNKNVALAANFLSRCPCSWANLVRMCDFTCSPOQSSPMKVVSTBEVDQ	119
Qy	143	-----AQLGAGOLPA--VVAYEAFYQHSFAEQSYDSYSCSRVRVPAAATLAVGTCGY	192
Db	120	EGESEKADATTAAPSPKEYITKIDIHITQYVINGTFPESCNCQSVSPSTQLDALDLCGDW	179

DT 01-MAR-2004 (Tremblrel. 26, Created)
 DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
 DE ENSANGP0000007352 (Fragment).
 GN ORFNames=ENSANG0000005547;
 OS Anopheles gambiae str. PST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 OC Anophelinae; Anopheles.
 OX NCBI_TaxID=180454;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RT "Anopheles gambiae re-annotation."
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=PEST;
 RG The Anopheles gambiae Sequence Committee;
 RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AA001008846; EAA06340.3; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD_5TM.
 DR Pfam; PF02460; Patched; 1.
 DR PROSITE; P850156; SSD; 1.
 FT NON_TER 1
 FT NON_TER 1003 1003
 SQ SEQUENCE 1003 AA; 111208 MW; P6B74BD2054D9E9A CRC64;

Query Match 25.5%; Score 1760; DB 2; Length 1003;
 Best Local Similarity 38.3%; Pred. No. 1.5e-115;
 Matches 402; Conservative 190; Mismatches 385; Indels 72; Gaps 23;

QY 243 CNEQGDVATSCDCAACPAIARPOALDSTFYLGQMPG---SLVLIILCSYPAVVT 299
 DB 1 CNEAY-DGYSACVDCDESCPTSEPPQPKDPGFVMVDLNGVTFVAVVVGIGLACIVL 59
 QY 300 ILLVGRVAPARDKSRVDPKGTSLDKLSP-----STHTLLGQFPQGWGTWASWPLT 354
 DB 60 ALLFG-----GREGSKQ-----QLPDLPSFGGPPSVNRALGRFTTWGTFCAENPVL 107
 QY 355 IIVLSVIPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILPA 414
 DB 108 ILAICSWIVGGLAFGIQYLIITDPVELWAAPDSRARQKDYFDSRFSPPYRTEQIFI-K 166
 QY 415 PNRSSVRYDSLLGPKNFGSILDLILLLELLOERLHLQWSPEAQRNLSLQDICYP 474
 DB 167 PTRQEFHHTAGNQTPGAYDREFLEVFKLQTTIE--QLGQAEGR---GLEQICFAP 221
 QY 475 LNP--DNTSLYDCINSLLYQFNRRNTLLLTANOT---LMGQTSQVDMKDFHLYCANAP 529
 DB 222 MTQVGAQTVLSECTQSVGFYFNS-----LAFENRTGDLNGFV--VNYLDKINGCTR-- 273
 QY 530 LTFKGTALALSCMADYGAVPFPFLAIGY-----KGKDYSEAEALIMTFSLNNYPAGD 583
 DB 274 -----NAYLPSCFGTYGGPIEPGVAVGGFPQAPGANDPRLATGVLTFLVEN-KANK 326
 QY 584 PRLAQKLEEAFLERAPQRMAGNFQVTTAERSLEDEINRTTAEDLPFATSYIVI 643
 DB 327 DELGPALEWEQRFVDMRDFEH---PMMEVAVSAERSIEDGDEMSEAEAMYTVIISYVM 383
 QY 644 FLVLSIALGYSWSRVMDSKATLGLGVAVVVLGAVMAAGFFSYLGRSSLVILQVVP 703
 DB 384 FVITFSLKINGRQPLHGRIVLAVGGIVVVLASVACSLGFFGVLEATMTMLTIEVIP 443
 QY 704 FLVLSVAGDNIPIFVLEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEACFFLGL 763

DB 444 FLVLAVGVNDIIFMLVHAFNRIDRERTPETACAIAGEALGQIGPSILITSECCFAIGAL 503
 QY 764 TMPAVRTEALTGLAVILDLQMSAFVALLSLDSKROBASRLDYCCVC---KPOBLPP 820
 DB 504 SMPAVNTFATVATVALLVDFLLQISAFVALMALDEQRVERGLDULVCCVRATNPKPV 563
 QY 821 PQQEGILLGFFQKAYAPFLHMTIRGVVLLFLALFGVSLXSMCHISVGLDQELALPKD 880
 DB 564 APDGPGLRVERFVVPVPELLRPRVLTVALFLVWGSLSLMVVPSEFGLDQELSMAD 623
 QY 881 SYLLDYFLFNRYFEVGAPVYFTTILGYNFSEAGNAICSSAGCNFFFTQKIQVATEF 940
 DB 624 SHVVKYFPEMAELFMWGPVYFVVRAGLNYTDVQHQLVCGGLTCLNDDSIETKLYQASLY 683
 QY 941 PQGSYLAI PASSWDDDFIDWLTPSSCCRLYISGPNKDKPCPTVNSLNCNKMCSITMGS 1000
 DB 684 PETHIARPASSWLDYIDWLAIQSCCKY---NPTDGSFCAS--NIVFPACPEEYDETG 738
 QY 1001 VRPSVEQPHKYLFWFLNDRPNIKCPKGLAAYSTSVN--LTS DGQ--VLASRFMAYHKPL 1056
 DB 739 IRPTVPQPERYLEFFFLSDLPDENCAKAGRAYSRALNYLLDRGLNVQDSYFSTVHTTA 798
 QY 1057 KNSQDYTEALRAARELANITADLRKVPCTDPAFVFPPTITNVTYEQYLTILPEGLFML 1116
 DB 799 VTSROFFYTALEQARLIAADIQMLDE---RQAGVEIFPYSVFVFEQYLTIVSDALQSL 855
 QY 1117 SLCLVPTPAVSCLLGLDLRSGLNLISVMTLVDTVGFMAWDISYNAVSLNLYSVAG 1176
 DB 856 GLSLAAVFTVFLVTGLDLSLVLLVLLVLMNMGLMNLNITLNAISLVNLVMSVG 915
 QY 1177 MSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQ 1236
 DB 916 IGVFEFISHIVRYLA-HGTRLESAAMVVRTGSSVFSGITLTKPAGIIVLAPAKSQIFQ 974
 QY 1237 IFFRLNLLITLGLLHGLVFLPVLSYV 1265
 DB 975 IFYFRMYLCIVLVAAGHLILLVFLTYI 1003

RESULT 24
 QSLNK7 MAGGR
 ID Q51NK7 MAGGR PRELIMINARY; PRT; 1275 AA.
 AC Q51NK7;
 DT 13-SEP-2005 (Tremblrel. 31, Created)
 DT 13-SEP-2005 (Tremblrel. 31, Last sequence update)
 DT 13-SEP-2005 (Tremblrel. 31, Last annotation update)
 DE Hypothetical protein.
 GN ORFNames=MG00819.4;
 OS Magnaporthe grisea 70-15.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
 OX NCBI_TaxID=242507;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=70-15;
 RA Birten B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
 RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
 RA Arachchi H., Armbruster J., Bachantsang P., Baldwin J., Barry A.,
 RA Bayul T., Blichsteyn B., Bloom T., Blye J., Boguslavskiy L.,
 RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
 RA Calvo S., Camarata J., Campo K., Chang J., Cheshateang Y., Citroen M.,
 RA Collymore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
 RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
 RA Dorjee K., Dorris L., Duffey N., Dupes A., Elkins T., Engels R.,
 RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
 RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
 RA Gnirke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
 RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
 RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Husby E., Iliev I.,
 RA Jaffe D., Jones C., Kamat M., Kamat A., Kamysseis M., Karlsson E.,
 RA Kells C., Kieu A., Kishner P., Kodira C., Kulbokas E., Labutti K.,
 RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,

RC STRAIN=FGSC A4;
RA Birren B., Nusbaum C., Abouelleil A., Allen N., Anderson S.,
RA Archchi H.M., Barna N., Bastien V., Bloom T., Boguslavskiy L.,
RA Boukhgalter B., Butler J., Calvo S.E., Camarata J., Chang J.,
RA Choepel Y., Collymore A., Cook A., Cooke P., Corum B., DeAtellano K.,
RA Diaz J.S., Dodge S., Dooley K., Dorris L., Elkins T., Engels R.,
RA Erickson J., Faro S., Ferreira P., Fitzgerald M., Gage D., Galagan J.,
RA Gardyna S., Gnerre S., Graham L., Grand-Pierre N., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Horton L., Hulme W., Iliev I.,
RA Jaffe D., Johnson R., Jones C., Kamal M., Kamat A., Karatas A.,
RA Kells C., Landers T., Levine R., Lindblad-Toh K., Liu G., Lui A.,
RA Ma L.-J., Mabbitt R., MacLean C., Macdonald P., Major J., Manning J.,
RA Matthews C., Mauceli E., McCarthy M., Meldrim J., Meneus L.,
RA Mihova T., Mienga V., Murphy T., Naylor J., Nguyen C., Nicol R.,
RA Nielsen C.B., Norbu C., O'Connor T., O'Donnell P., O'Neill D.,
RA Oliver J., Peterson K., Phunkhang P., Pierre N., Purcell S.,
RA Rachupka A., Ramasamy U., Raymond C., Retta R., Rise C., Rogov P.,
RA Roman J., Schauer S., Schupback R., Seaman S., Severy P., Smirnov S.,
RA Smith C., Spencer B., Stange-Thomann N., Stojanovic N., Stubbs M.,
RA Talamas J., Tesfaye S., Theodore J., Topham K., Travers M.,
RA Vassiliev H., Venkataraman V.S., Viel R., Vo A., Wang S., Wilson B.,
RA Wu X., Wyman D., Young G., Zainoun J., Zembek L., Zimmer A., Zody M.,
RA Lander E.;
RT "Genome Sequence of Aspergillus nidulans";
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AACD0100032; EAA64951.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1271 AA; 139921 MW; 58ADD27F16364A74 CRC64;

Query Match 24.6%; Score 1700.5; DB 2; Length 1271;
Best Local Similarity 31.7%; Pred. No. 3.4e-111;
Matches 424; Conservative 249; Mismatches 493; Indels 171; Gaps 41;

QY 25 TTIHQGYCAFVDEGCKNPESLSMTLSNVSLNTPARKITGDHLLILQKIC-PRLYT 83
DB 25 TRIHENGRCIAIRGHCKQKQFFGEL-----PCPDNDAREPEAAVREKLVNLCQAKQWE 78

QY 84 GNTQACCSAKQLVSLSEASLITKALLTRCPACSDNFVNLHCHNTCSNPQSLFINVTRVA 143
DB 79 GP---VCCBEEDIDALSKNLKAEGIISSCPACKDNFFNICTFCSPDQSLFINVTKTE 135

QY 144 QLGAGQLPVAVEAPYQHSFAEQSDSCSRVRVPAATLAVGTWCVGVSALCNAQRWL 203
DB 136 KNGSGK-ELVTELDNIWSBEYOSGFYDSCNKVNGKSGGKAIDFI----GGGAKDYQOPL 190

QY 204 NFGDGTNGLAPLDITFHLLPQAVGSGIQPLNEGVARCNSQGDVATCSQDCQCAASC 263
DB 191 KFLGDKKFLGSPFQINHTPEPDSQGMALFHP--KACNDA--DPAYRCSVDPCDVC 246

QY 264 PAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTILLVGF-----RVA 308
DB 247 PELPAIKT-EHCHVGLLPCLSFVILIYSVF---LLGVAGFSYSFTYRERYRPERVR 302

QY 309 PARDKSNWDPKKG--TSLSDKLS-----FSTHTLQGFQGWCTVWASWPLTILVLSVI 361
DB 303 LLQDPNPSDEDEGDIVFHAGGHLEYPHGYKLNLSMLDTVFSRIGSVYCARFPALTIISSVV 362

QY 362 PVVALAAGLVFTETLTPVVELWSPAPNSQARSEKAFHDQFPPFRNQVILTAPEPNSYR 421
DB 363 AVLLSLGWLWURFVETDPVRLWVSPISAAKEKAFDENFGPPYRAEQAFVNDDET--- 419

QY 422 YDSILLGPKNFSGILDLLELLELQELRHILQVMSPEAQRNISLODICYAPLNPNTS 481
DB 420 -----GDGRVLDYDTLTWTFGVESIRRV-----ISLDRLSLDDICYKPTGD--- 462

QY 482 LYDCCINSLLOYQNNRTLLLTANQTLMTQTSQVD---WKDFLCANAPLTFKDGTL 538
DB 463 --ACVIOQSVTYGFG-----GSLNLDPTWQDRLTHCASSP-----GDA- 499

QY 539 ALSCMADYAGVPFPFLAIGGYKKG-DYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFL 597

DB 500 --SCLPDFSQPLRPEMILGYEDSGNVLDKALIVTWVNNHAPGSEEAEDWEDTFR 557
QY 598 EEMRAFORRMAGM-FQVTTAERSLEDEINRTAEDLPFATSVIVILYISIALSGYS- 655
DB 558 GIFQVQEEAKNRGLRVSTFTEASVQELNSNTDAKIVIVISIIIMFIYASIALGVSVM 617
QY 656 SW-----SRVWVDSKATLGLGVAVLGAVNAAMGFFSYLGRSLVILQVVPFLVLS 708
DB 618 TWASLINNPANALVQSKFTLGVGVIVVLMVSASVGLFSAAAGVKVTLIAEVIPFLVLA 677
QY 709 VGADNIFIFVLEYQRL-PRPGEPRREVHIGRALGRVAPSMMLCSLSEAI CFFGALT PMP 767
DB 678 VGVDNIFLIYVEPERLVNHPDEIDERISRAIGRIGPSIFLSAITETVAFALGVFVGMP 737
QY 768 AVTFALTSGVAILVDFLOWSAFVALLSDSKQERASRLDVCCV-----KPOELP 819
DB 738 AVRFNFIYAAGAVFINAVLQITMFSVIALNQKRVESLRADICPCLTVKRAHSGMPEDLA 797
QY 820 PPGQG-EGLLGFFOKAYAPFLHWTGRVLLFLFALFGVLSYSMCHI SVGLDQELALP 878
DB 798 FDDQREGIILQKIRKYAPALLNRRVKVVVITFLGILAAGLALTEVAMGLDQRIALP 857
QY 879 KDSYLLDYFLFLNRYFEVGAPVYVVTTLGVNFSEAGMNAICSS-AGCNNFSPFTQKIQA 937
DB 858 SDSYLLIDYFDLSEYFNSGPPVYFVTR-NVNITKREHORQLCGRFTTCBEYSPLFVLEQE 916
QY 938 TEPPEQSYLAIPASSWDDFDIMLTP--SSCRLYISGPNKDKFC-----PSTVNSLNLCK 991
DB 917 SKRNSVSIAGATASWIDDFYMLNPOODCC--YEDG----KLCFEGRTGWNISLTGM- 969
QY 992 NCMSITMGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLAAAYTSV-----NLTSDGQVLA 1046
DB 970 -----PGAEFIHYLEKWKISPTDASCPLGGKAPYSNALVDPDKRITTN----A 1014

QY 1047 SRFWAYHKPKNSODYTEALRAARELANITADLRKVPGTDPAREVEPYTITNVEYQYL 1106
DB 1015 SHFTSTHTPLRTQDDFIKSVISARRIADGLSA-----EHGIDVFPYSKTYIFFDQYV 1066

QY 1107 TILPEGLFMLSCLVPTFAVSCLLGLDLRSLNLNLLSVMLVDTVTGFMALWDISVNAV 1165
DB 1067 SIVQAGTTLGSAVAIIFAITSIILG-SVATGAVVTATVIMVIDIIGSMAISGVSLNAV 1125

QY 1167 SLINLVSAGVMSVEFVSHITRSPAISTK-----PTWLE-----RAKEATISMGSAVFAGV 1216
DB 1126 SLVNLVICVGIGVEFCAHIAAFWPPSKRIMEIVPSKFRGDKARSWTALVNVGSGVFSGI 1185

QY 1217 AMTNLPGLIVLGLAKAQLIQIFPRLNLLITLLGLLHCLVPLVLTLSVVGPD--VNPALA 1274
DB 1186 TVTKLLGCVLAFTRSKIFEIYFRVWLALIIFAATHALIFLPVALSYFGDGVADPEAS 1245

QY 1275 --LEQKRAEEAAVAVMV 1289
DB 1246 GGLEENLASRSTRSLIV 1262

RESULT 27
QSKGS9_CRYNE PRELIMINARY; PRT; 1330 AA.
AC QSKGS9;
DT 10-MAY-2005 (TremBLrel. 30, Created)
DT 10-MAY-2005 (TremBLrel. 30, Last sequence update)
DT 10-MAY-2005 (TremBLrel. 30, Last annotation update)
DE Vacuolar membrane protein, putative.
GN ORFNames=CNE02470;
OS Cryptococcus neoformans var. neoformans JEC21.
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
OX NCBI_TaxID=214684;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=JEC21;
RA Loftus B., Amedeo P., Roncaglia P., Vamathevan J., Utterback T.,

RA Van Aken S., Fraser C.;
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
 RC STRAIN=JEC21;
 RX PubMed=15653466; DOI=10.1126/science.1103773;
 RA Loftus B.J., Fung E., Roncaglia P., Rowley D., Amedeo P., Bruno D.,
 RA Vamathevan J., Miranda M., Anderson I.J., Fraser J.A., Allen J.E.,
 RA Bosdet I.E., Brent M.R., Chiu R., Doering T.L., Donlin M.J.,
 RA D'Souza C.A., Fox D.S., Grinberg V., Fu J., Fukushima M., Haas B.J.,
 RA Huang J.C., Janbon G., Jones S.J.M., Koo H.L., Krzywinska M.I.,
 RA Kwon-Chung K.J., Lengeler K.B., Maiti R., Marra M.A., Marra R.E.,
 RA Mathewson C.A., Mitchell T.G., Perteau M., Riggs P.R., Salzberg S.L.,
 RA Schein J.E., Shvartsbeyn A., Shin H., Shumway M., Specht C.A.,
 RA Wye N.H., Tenney A., Utterback T.R., Wickes B.L., Wortman J.R.,
 RA Wye N.H., Kronstad J.W., Lodge J.K., Heitman J., Davis R.W.,
 RA Fraser C.M., Hyman R.W.;
 RT "The genome of the basidiomycetous yeast and human pathogen
 Cryptococcus neoformans";
 RL Science 307:1321-1324 (2005).
 DR EMBL; AB017345; AAW43643.1; -; Genomic_DNA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD_5TH.
 DR Pfam; PF02460; Patched; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Complete proteome.
 SQ SEQUENCE 1330 AA; 146237 MW; 223B9F89F35EB5A5 CRC64;

Query Match 24.0%; Score 1659.5; DB 2; Length 1330;
 Best Local Similarity 30.6%; Pred. No. 2.9e-108;
 Matches 427; Conservative 247; Mismatches 448; Indels 275; Gaps 40;

QY 31 GYCAFVDEGKPNELSGSLMTLSNVCLSNTARKITGDHLLLOKICPRLYTGNTQAC 90
 DB 11 GICMRGTGCK-----TSMGADLPDSDSDATVPDQRLDLMAVCGPSYSLPD-HVC 63

QY 91 CSAKQLVLEASLSITKALLTRCPACSNFVNLHCHNTCSNQSLFINVTRVAQAGAGOL 150
 DB 64 CTYDQLSTLSDRLQAAPLIASCACINNFRSFCDFTCSPNQSTFLSVATQKTEGK- 122

QY 151 PAVAYEAFYQHSFPAQSDSCSRVRVPAATLAVGTGCVGVSALCNAQRLNFGQDTG 210
 DB 123 DAVKEVDYEVSSDFKGFYDSCQVQFATNGFAM-----DLTGGCATNASGFLKTMGDLR 178

QY 211 NGL-APLDTITLHLLPQAVSGIOLNEGVA-----RCNESQDDVATCSCQCAASC 263
 DB 179 PGLGSPFQINF-----PDNDDSAYRRAPLSCSDAENIN-ARCAADCPSCVC 223

QY 264 PA---IARPOALDSTFYLGQMPGSLVLIILCSVFAVVTILVGPVAPARDKSK----- 315
 DB 224 PSLPIAPST--KQCHGVANSLTFLSIITYSVILVGLLYIWKQA-AHRQRBYERV 280

QY 316 -MVDPK-----KGTSLSDKLSF-----S 332
 DB 281 ALLDPHSPPTIQNGQGLDGLMGRNDDAEGSPGSIHFRLGRGASLLDPMEHLQPKNQK 340

QY 333 THTLQGFQFGWTVASWPLTILVSVIPVALAAGLVFTLTDPVELWSAPNSQARS 392
 DB 341 INATLRRFFYRLGLTCAKRIEVEFAISALI VGLFNFGWKYFEVEVDPPVRLMVSPTSAS 400

QY 393 EKAFHDFQFGPPFRNTQVILTAAPNRSSVRYDLSLLGPKNFSGILDLLELLELLEQLRL 452
 DB 401 QKQFDDSDGFFYKSEQIFITQSSGSPVNYDT-----LDWLKLV----- 439

QY 453 HLQWNSPEAQRNISLQDICYAPLNPNTSLYDCCINSLLQVFNQNRLLTLLTANQTLMGQ 512
 DB 440 ETEINALKTSQDGLGLEDICFAPAGKGT---PCVIOVSVAW-----LGD 479

QY 513 TSQV---DWKHFLYCANAPLTFKDGATALASCMDYCAPVFPFLAIGYKGYKDYSEABA 569
 DB 480 DMENVGEEWESRVRDCASRP-----GECLEPPFGQPIDPKLVLGANG-DWLEAKA 528

QY 570 LIMTSLNNYPAGDPRLAQAKLWEBAFLSEMAFQRMAGMAGFQVTFTAERSLEDEINRTT 629
 DB 529 LVVTVWVWNNY--NDERVEPAEQWERKLDYLDLKR--PGI-KISVSTGVSEEBINKST 583

QY 630 AEDLPIFATSYIVIFLYISLALG-----SYSSWSRV----- 660
 DB 584 NTDVKLVVLSVLMFFYVSLTILGGGLPPSMIOAFTHRAYRLVFRIGLVLLHVKDAPLES 643

QY 661 -----MVDSKATIGLIGGAVVVLGAVMAAMGFSSVYLGRSSLSVLQVVPFLVL 707
 DB 644 APPPDFRIVPTLLSVNKFESLGLFGIVIVLIVAVSSVGLFSLGVRVTLIIAEVLPFLVL 703

QY 708 SVGANINIFVLEYOR-----LPRRPGPREVH-----IGRALGRVAPS 746
 DB 704 AVGVNDFILVHELEQRNNLHAAQCGDDES VHSNGAQPSTFLAPEBRVARAVARMGFS 763

QY 747 MLLCSLSAICFFIGALTFMPAVRTFALTSLGLAVLDFLLOMSAFVALLSLDSKQESAR 806
 DB 764 IMLSSVTEVAVFALGALVPMFAVRNFAIYAGSVLFGAIMOCTVFVSAMTDLRSESME 823

QY 807 LDVCCCVKPOELPPPG-----QCEGLLGFQKAYAPFLHHTITRGVVLLLELALFCV 859
 DB 824 IDCFCIRLR--PPGLYDNEAPREGWVKFMRVTYAPSLRHEVHKQVLVAFGLFLA 881

QY 860 SLYSMCHISVGLDQELALPKDSYLDYFLFLNRYFEVGAPVYFVTTLGYNFSEAGMNAI 919
 DB 882 AIIQIHISLGLDQRLALPSSSHLVYPVNDVSDYLDVGPVYFVTE--GDPSPSRHQOQL 940

QY 920 CSS-AGCNFSFTQKIQYATEFPPEQSYLAI PASSWDDPIDMLTPS--SCCR-----L 969
 DB 941 CGREFTTCLSESVANTLEAERKRPDSSTASPPAAWIDDFLQWNTNPTFSCCRVRDRPTI 1000

QY 970 YISGNKDKFC-----PSTVNSLNCNCHSITWGSVRSVPEQFHKLPWFLNDRP 1020
 DB 1001 FCSPKDSERLCRCPCEGQKWDSTMDGL-----PEGEDFMRYLEQWLSIPT 1045

QY 1021 NIKCPKGGLAAAYSTSVNL--TSDGOVLAERFMAYHKPLKNSQDYETALRAARLAANITAD 1079
 DB 1046 NDECPGQQAQPYSAVKLVSSNTTVAASHFRYTHYTPKASQADFINALAAARRISDIT-- 1103

QY 1080 LRKPGTDPAPVEPPYTTINVPYQYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRSL 1139
 DB 1104 -----HRTGVRVPYSLFYFFQYEHIIISMAIQVLFLAVLVITSTLIG-SWRITGG 1156

QY 1140 LNLISIVMLVDTVGFPMALWDISYNAVSLINLVSAGMSVRSVSHITSP----- 1189
 DB 1157 TVTFTCALAVINVGMVGYWGSLNALSINLVLSLGLIAGVAFCSHARAFMGAGSLPLD 1216

QY 1190 AISTKPTWLERAKATISMGSAVFAVAVANTMLPGILVLGLAKAQLIQIPFFRLNLLITLL 1249
 DB 1217 KVEGKERDERAWTALVDVGPSPVSGITWTKLIGISVLALTSKLELVYFFRMWLSLILS 1276

QY 1250 GLHLGLPVLPSVYG 1266
 DB 1277 GALHGLVLLPVLLSYLG 1293

RESULT 28

Q55SD4 CRYNE

ID Q55SD4 CRYNE PRELIMINARY; PRT; 1330 AA.

AC Q55SD4;

DT 13-SEP-2005 (TremBLrel. 31, Created)

DT 13-SEP-2005 (TremBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TremBLrel. 31, Last annotation update)

DE Hypothetical protein.

GN ORFNames=CNB2480;

OS Cryptococcus neoformans var. neoformans B-3501A.

OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;

OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.

OX NCBI_TaxID=283643;

RN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=B-3501A;
RA Fung E., Hyman R.W., Rowley D., Bruno D., Miranda M., Fukushima M.,
RA Wickes B.L., Fu J., Davis R.W.;
RT "Cryptococcus neoformans serotype D sequencing";
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA01000024; EAL20887.1; -; Genomic_DNA.
KW Hypothetical protein.
SQ SEQUENCE 1330 AA; 146208 MW; 77DD1023070B520A CRC64;

Query Match 24.0%; Score 1658.5; DB 2; Length 1330;
Best Local Similarity 30.6%; Pred. No. 3.5e-108;
Matches 427; Conservative 246; Mismatches 449; Indels 275; Gaps 40;

QY 31 GYCAFYDCGKNPELSGSLMTLSNVCLSNTPARKITGDHLLQKICPLRYTGPNTQAC 90
DB 11 GICAMRGTCGK-----TSMFGADLPCDDSDATVPDOSLLDMSVCGPSYSLPD-HVC 63

QY 91 CSAKQVLSLEASLSITKALLTRCPACSDNFVNLHCHNTCSNQSLFINVTRVAOLGAGQL 150
DB 64 CTYDQLSTLSDRLOQAAPLIASCPACINFRSFYCDFTCSNQSTFLSVTATOKTTEGK- 122

QY 151 PAVAYEAFYQHSFAEQSYDCSRVRVPAATLAIVTGWGVGSALCNAQRWLNFGQDTG 210
DB 123 DAVKEVDYEVSSDFKQGFYDSCDKVQFGATNGFAM-----DLTGGGATNASGLKTYMGDLR 178

QY 211 NGL-APLDTTHLLPQAVGSGIOPLNEGVA-----RCNESQDDVATCSQDCAASC 263
DB 179 PGLSGPFFQINF-----PDNDDSYRREAPLSCSDAENIN-ARCAACDCPSVC 223

QY 264 PA---IARQALDSTFYLGQSGSLVLIILCSVFAVYTLVGRVPAPARDKSK----- 315
DB 224 PSLPIYAPST--KQCHGAVSCLTFLSLIIYSVIILVYKQALYIWKQA-ARHRORRYERV 280

QY 316 -MVDEK-----KQTSLSQKLSF-----S 332
DB 281 ALLDDPHSPSTIQNGQSGDGLGMRNDDAESGPGSIHFLRGSGASLLDPMELHQPKNK 340

QY 333 THTLGQFFQGGTGWASWPLTILVLSVPVVALAAGLVFTLTTPDVELMSAPNSQARS 392
DB 341 INATLRRFYRLGLTCAKPEIEVFALISALVGLFNGWKYFEVETDPLVWLWSPTSAS 400

QY 393 EKAFHDHGFPPFRTNQVTLTAPNRSSRYSDLSLLGPKNFGSLDLDLLELLELQERL 452
DB 401 QKQFFDSTFGPFSQIFITQSSGSPVNYDT-----LDWWLKV----- 439

QY 453 HLQVMSPEAQRNLSLODICVAPLNPNTSILYDCINSLILQYFQNNRTLLLTANQTLMGQ 512
DB 440 ETEINALKTSDDGIGLEDICFAPAGKGT---PCVIGSVSAW-----LGD 479

QY 513 TSQV---DWKDFLYCANAPLTPKDGTALALSQADYGAIPFPFLLAIGGYKGYSEABA 569
DB 480 DMEVWGEESRVRDCAAP-----GECLEPPFGQPIDPKLVLGANG-DWLEAKA 528

QY 570 LINTFSLNNYPAGDPRIQAOKLMEAFLEBMAFQRMAGMFGVTTAERSLEENRT 629
DB 529 LVVTVWVNNY--NDERVEPAQWERKLDYRLDKR--FCI-KISYSTGVSLEEEINKST 583

QY 630 AEDLPFATSVIIVLYISIALG-----SYSSWSRV----- 660
DB 584 NTDVKIVLSYLVNMFVVSUTLGGGLPPSMIOAFTHRAVRLVFRIGVLLHLVKDAPLES 643

QY 661 -----MVDSKATILGGVAVVLGVAAMAGFFSYLGIRSLVILQVFPFLV 707
DB 644 APPDPFRVITLLSVNSKPSLGLFVIVILVIAVSSVGLPSLLGVRVTLIAEIVPFLV 703

QY 708 SVGADNIFIFVLEQV-----LPRPGEPREHV-----IGRALGVAPS 746
DB 704 AVGVNDVFIILVHELRQNNLHAAQQPGDDSVHSGNAQPSGTTFLAEEVARAVARWGPS 763

QY 747 MLLCSLSEATCFFLGALTTPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKROEASR 806

DB 764 IMLSSVTEVVAFGALVEMPAVRNFAIYAAGSVLFGALMQCTVFSAMTLDLRRSESMR 823
QY 807 LDVCCCVCKQOELEPPGP-----QGBGLLGFQKAYAPFLLHWTTRGVVLLFLALFGV 859
DB 824 IDCFCPCIRLR--PPIGLYDNEAPSRGEMVKFMRTVYAPSLLRHREVKQLVAVFGGLFLA 881
QY 860 SLYSMCHISVGLDOELALPKDSVLLDYFLFLNRYPEVGPAPVVPVVTTLGVNFSSEAGNAI 919
DB 882 AIIQIOHISLGLDQRLALFSEHLYVEYFNDVDSYLDVGPVPVTE--GGDPSRHGOQL 940
QY 920 CSS-AGCNFSPFKIQYATEPEEQSYLAIPAASSWVDDFDMLTPS--SCCR-----L 969
DB 941 CGRFTTCLLSVANTLEAEKRPRDSSFIASPPAAWIDDFLOWTNPNTFESCCRRRRDPTI 1000
QY 970 YISGPNKDKFC-----PSTVNSINCLKNCMSITMGSVRSVPSEQFHKYLPWFLNDRP 1020
DB 1001 FCSFKDSIRCRPCFEGQKWDSTMDGL-----PEGEDFMXYLSQWLISPT 1045
QY 1021 NIKCPKGGLAAYSTVNL--TSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITAD 1079
DB 1046 NDECPGQGPAPYAAVKLVSSNTTVAASHPTVYHPLKSOADFINALAAARRISDDIT-- 1103
QY 1080 LRKVPGTDPAPFPFPTITNVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLSGL 1139
DB 1104 -----HRTGVRVFPYSLSFYVFPDQYEHIIISMAIQVLFVAVLVITSTLLG-SMRTGG 1156
QY 1140 LNLISVILVDTVGFMALWDISVNAVSLINLVSAGVMSVEFVSHITRSP----- 1189
DB 1157 TVTFTCALAVINWGVNGYGISLNAISLVNLVLSLGIAVEFCSHARAPAGAGSGLPD 1216
QY 1190 AISTKPTWLRRAKATISMGSAPVAGVAMTNLPGLVGLAKAQLOIFFRMLNLLITLL 1249
DB 1217 KLEGRKEDERAWTALVDVGPSPVSGITMTKLGISVLALTRSKLSEVFRMWLSLILS 1276
QY 1250 GLLHGLVPLPVILSYVG 1266
DB 1277 GALHGLVLLPVLLSYLG 1293

RESULT 29
Q4WNG5 ASPFU PRELIMINARY; PRT; 1273 AA.
ID Q4WNG5
AC Q4WNG5
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Patched sphingolipid transporter (Ncr1), putative.
GN ORFNames=Afu609980;
OS Aspergillus fumigatus Af293.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=330879;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Af293;
RA Nierman W., Pain A., Anderson M.J., Wortman J., Kim H. Stanley.,
RA Arroya J., Berriman M., Abe K., Archer D.B., Bermejo C., Bennett J.,
RA Bowyer P., Chen D., Collins M., Coulson R., Davies R., Dyer P.S.,
RA Farman M., Fedorova N., Fedorova N., Feldblyum T.V., Fischer R.,
RA Fosker N., Fraser A., Garcia J.L., Garcia M.J., Gobie A.,
RA Goldman G.H., Goni K., Griffith-Jones S., Gwilliam R., Jimenez J.,
RA Haas H., Harris D., Horluchi H., Huang J., Humphrey S., Jimeenez J.,
RA Keller N., Khouri H., Kitamoto K., Kobayashi T., Kulkarni R.,
RA Kumagai T., Lafton A., Latge J.-P., Li W., Lord A., Lu C.,
RA Majoros W.H., May G.S., Miller B.L., Mohamoud Y., Molina M., Monod M.,
RA Mounia I., Mulligan S., Murphy L., O'Neill S., Paulsen I.,
RA Penava M.A., Perteau M., Price C., Pritchard B.H., Quail M.A.,
RA Rabinowitz E., Rawlin N., Rajandream M.-A., Reichard U.,
RA Renauld H., Robson G.D., Rodriguez de Cordoba S., Rodriguez-Pena J.M.,
RA Ronning C.M., Rutter S., Salzberg S.L., Sanchez M.,
RA Sanchez-Ferrero J.C., Saunders D., Seeger K., Squares R., Squares S.,
RA Takeuchi M., Tekala F., Turner G., Vazquez de Aldana C.R., Weidman J.,

DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.0601; 1.
DR TIGRFAMs; TIGR00917; 2A060601; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1275 AA; 141153 MW; 375FCE96BED2C9FF CRC64;

Query Match 22.1%; Score 1527.5; DB 2; Length 1275;
Best Local Similarity 29.4%; Pred. No. 6.4e-99;
Matches 413; Conservative 223; Mismatches 419; Indels 349; Gaps 47;

QY 31 GYCAFDYDCGKNPELSGSLMTLSNCLSNTPARKITGDHLI--LLQKICRLTYGPNQ 88
DB 17 GYCAMYDLCGARS--GKVL-----NCPFNIPS--VKPDDLSSKIQSLCPTI-TG---N 63
QY 89 ACCSAKQVLSEASISUTKALLTRCPACSDNFVNLHCNTCPNOSLPINTRVAQLGAG 148
DB 64 VCCTETQPTDLRSQVQAIPITVGCPCACLRNFLNLCETLCPDQSLFINVTSTTKVKN 123
QY 149 QLPVAVVEAFQHSFAQSVDSCSRVPAATLAVGTMCGVGSALCAQRLNFOGD 208
DB 124 STVDGIQY--YTDDFGAGMYESCKNVKFGSNSRALDPL-----GAGAKNFKWFTFIQ 177
QY 209 -TGNGL--APLDITFLLPEGOAVSGIQPLNEGVARCNESQGDVATCSQDC--AASC 263
DB 178 KAGVNLPGSPYGIAP--LPTPPVSSGMRPMNVSIYSC---GDESIGCSGDCPCSAATC 230
QY 264 PA--IAAPQALDSTFYLGOMPGSLVLIILCSFVAVTILLVGRVAPAPARDKSK----- 315
DB 231 SSKAEVPTQKHSCSIKIGSEVKVCDFILATLYLVLSFLGGGLHPVRGKKTQSMG 290
QY 316 -----WDPKKGTSLSKLFSFT-----HTLGGQFGMGWTVASWP 352
DB 291 TLSEASGERNSVNOQRKPTIQSQMLQNTQPNRWGQLSTVOGH--LANFYKGIWVARHP 348
QY 353 LTILVLSVPPVALAGLVFTLTDTPVELNSAPNSQARSEKAFDQHPGFPFRNTQVIL 412
DB 349 TLVLCLSVSVLLLCVGLIRFKVETRPDKLWVGSGSRAAEKQFDTLHAPPYRIEQLII 408
QY 413 TAPNRSSVRYDSLLGPKNFGSILDLLELLELQERLHLQVSPSPAQRNLSIDICY 472
DB 409 ATVOTSSHEKAPEILTDNDIKLLFDI-----QKKVS--QLFS 443
QY 473 APIN--PDNTSLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDM--KDHFLYCANAP 529
DB 444 NPSNHPYVFMRTCKGLFNMIFK-----MKPENYDDYGGVDHVKYC----- 485
QY 530 LTFKQGTALALSCMADYGAVPFPFLAIGYKGYKDYSEAEALIMTFLSNYPAG--DPLRA 588
DB 486 --FEHFTSTE--SCLSAFKGLDPTTALGGFGSGNSPSEASAFVTVVDNFVNDKGNKTEK 542
QY 589 AKLWEAEFL--EMRAFQRMAGMFQVTTAERSLEDEINRTAEDLPFAT----- 638
DB 543 AVAMEKAFIQLAKDELPLPMVOAKNLTLSFSSESSIEELKELSTADVITIAVICFSFILY 602
QY 639 -----SYIVIPYISLALGSYSSRSKRVMDSKATGLGGSVAVLGA 679
DB 603 WTSNMGFMSISHVSLQISLVWFAYISLTGDSPLKSFYITTSKVLGLGSLVLLVMS 662
QY 680 VMAAMGFFSYGLRSSLVLQVPPFLVS--VGADNI-----PI 716
DB 663 VLGSVGFSAVGMSKTLIIMEVIFPLVLAIVISISNIACNFNMLADVATFTFLIFIFY 722
QY 717 FVLEY-----QRLPRRGE--PREVHIGRALGRVAPSMLLCSLEACFFLG 761
DB 723 FYLEYFYRGVQVDMCMILVHAVRQEQELPERRISNALMEVGPSTLASLAELAFAG 782
QY 762 ALTPMPAVRTFA-----LTS-----GLAVILDPLQMSAFVALLSDSKR 801
DB 783 APTKMPAVRVFSPFVAVSNFAFIYFETSCIMLAALVLLDQLQITAFVALIVDFRR 842
QY 802 QEASRLDVCCVKPQE--LPPPGQGE--GLLGGFFQKAYAPFLHWTIRGVVLLP-- 853

843 TDKRVDCFPICIKTSKSSISAEGVGQKAGLTRYMKVHAPVLSHWIKVIVIAFFG 902
QY 854 LALFGLSVLSWCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPYFVFTTLGYNFSS 913
DB 903 LAMAGIALST--RIEPOLEQIVLPQDSYL----- 930
QY 914 AGMNAICSSAGCANNFSTQKIQVATEPEPESYLAIPASSWDDPIDWLTSS--CCRLYI 971
DB 931 -----QIARASLTPELSYIAKPAASWLDLDFVLSPEAFGCCRKFT 971
QY 972 SGNPKDFCSTVNSLNLKNCMSITWGSVRPSVEQPHKVLFWLNDNRNPKCPKGGGLAA 1031
DB 972 NG-----TFCPP-----DLSSDRPSTQPKKELFWLNLALPSADCAKGGHGA 1021
QY 1032 YSTSNNLT--SDGOVLASRFMAVHKLNQSDYTEALRAARELANITADLRKVPGTDP 1089
DB 1022 YSSVDLQGVANGIIQASSFTYHPLNKQVDFVNSRAAQEFSAKVSRLK----- 1073
QY 1090 FEVPYITTVVBYQYLTIPEGLFMLSLCLVPTFAVSCLLGLDLRLSGLLNLSVIMIL 1149
DB 1074 MEIYPYSVFMFPQYLDIWKTLINLSIAIV----- 1105
QY 1150 VDTGFMALWDISVNAVSLINLSAVGMSVEFVSHITRSPASTKPTWL----- 1198
DB 1106 -----PHIQNLALSVNLINSVGLAEFCFVHIHAFSVSLSVCTSIHKKQFAIVT 1156
QY 1199 -----ERAKEATISMGSAFAGVAMTNLPGLILVGLAKAQLIPIFFRLNL 1244
DB 1157 LTVRNAQISTGDRNHRMKEALGGMASVFSGITLTKLVGVIVLGSFSEVFWYFKMWL 1216
QY 1245 LITLLGLLHGLVLPV---ILSYV 1265
DB 1217 ALVLLGLHGLVLPVSHGVLEHV 1240

RESULT 33
Q9TVK6_DICD1
ID Q9TVK6_DICD1 PRELIMINARY; PRT; 1342 AA.
AC Q9TVK6
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)
DE NPC1 protein.
GN Name=npca; ORFNames=DD80191127;
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX3;
RA Xu J., Postlethwait P., Steck T.L.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AX4;
RA Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA Sugeng B., Berriman M., Song J., Olsen R., Szafarski K., Xu Q.,
RA Tunggal R., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,
RA Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA Muzzy D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA Louiseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA Urushihara H., Hernandez J., Rabbintovitch E., Steffen D., Sanders M.,
RA Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA Shauleky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,

CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

DR EMBL: AABX01000767; BAA26850.1; -; Genomic DNA.

DR GO: GO:0016021; C: integral to membrane; IEA.

DR GO: GO:0016020; C: membrane; IEA.

DR GO: GO:0008158; F: hedgehog receptor activity; IEA.

DR InterPro: IPR000152; Asx hydrosyl S.

DR InterPro: IPR003392; Patched.

DR InterPro: IPR00731; SSD 5TM.

DR Pfam: PF02460; Patched; I.

DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.

DR PROSITE: PS0156; SSD; I.

KW Hypothetical protein.

SQ SEQUENCE 1162 AA; 127777 MW; 6776A78BB9B9619 CRC64;

Query Match

Best Local Similarity 21.7%; Score 1500; DB 2; Length 1162;

Matches 383; Conservative 229; Mismatches 437; Indels 170; Gaps 35;

QY 123 LCHHTCPNQSLFNVTRVAQAGALPQAVVAYEAFYQHSFAEQSYDSCSRVRVPAAT 182

DB 1 MECTTCSNQSLFNVVTKTIEKKGL--VTELDQLISEEYGTGFYNSCKDVKGPTNS 58

QY 183 LAVGTWCGYGGALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGGIQLNEGVAR 242

DB 59 RAM----DLIGGAKNYTQLLKFLGQERFGSGPFOINF----PVEYAPDMKPLPMKPKK 110

QY 243 CNEOGDDVATCSODCAACPAIARPOALDSTFFYLGOMPGSLVLIILCSPFAVVTILL 302

DB 111 CNDE--DPNFRCAVDCPEICPTLPDVEQAGC-HVGAFLCPLSFASILTYSVILFISAA 167

QY 303 VGRFV-----APARDKSK-----MVD-PKKGTSLSKLSFSTH 334

DB 168 VGVHVAKEHAKRSERLLELTDAPEDEDEGDTQNVAMIDRQK-----TYIN 219

QY 335 TLGGQFFQCGWGTWASWPLTILVSVIPVVALAAGLVFTETLTPDVELWASPNQARSEK 394

DB 220 TWCDSAFSLKGVAAFPFAITVTSILIASILSLGWFFHELEKPNARLWVSPSPAEEK 279

QY 395 AFHDHGFQFFRTNQVILTAPNRSYRVDLSLLGPKNFGIILDLLELLELQRLRHL 454

DB 280 AFFDSHFHGFAYAEKVFELVNDTPQS-----GP---GPVLSRDLTLLWMDVEKSVAA 328

QY 455 QVMSPEAQNRISLQDICVAPLNPNDTSLYDCCINSILOVFNQNRLLLTANQTLWGQTS 514

DB 329 K-----GSNYGSFQDLCKPTGD-----ACVQSVAAVFDQD-----PD 363

QY 515 QVD---WKDHFLYCANAPLTFKQGTALALSCMADYGAPVFPFLAIGGY-KGKDYSEBAL 570

DB 364 SVDPTWQSTLTCAASP-----VECRPAYGQPLDPSMILGGYPEGGNVAEASAM 413

QY 571 TMTFSLNNYPAGDPPLAOKLWEAFLEBMAFQ--RRWAGMPQVTPFAERSLEDEINRT 628

DB 414 TTVTWLINPSENSPEVDAMDWEVALKRLLEVEQDEAKERGL-RLSFSTETSLSEELNKS 472

QY 629 TAEDLPPIATSVIVIFLYISLALGSYSSNSR-----VMYDSKATLGLGGVAVVLGAV 680

DB 473 TWTDAKIIVISIIIFLYASLAGSTLTLPKDLINPVAHSVLEKFTLGIVGIVILMSI 532

QY 681 MAAMGFFSYGLSRSLVLQVVPFLVSLVGADNIFIFVLEYQL-PRRPGEPREYHIGRA 739

DB 533 TASIGLFWAGLRATLIIVDVIPFVLAVGVNDIFLIVHEPERVNVSPDDMVBARISRA 592

QY 740 LGRVAPSMLLCSLSBAICFFGLATFMPAVRTFALTSGLAVILDFLLQMSAFVALLSDS 799

DB 593 LGRMGPSILFSALTATSPALGAFVGMFAVRNFAIYAAGVAFINAILQVTFEVSVLTLNQ 652

QY 800 KQEASRLDVCCKVQKQ-----ELPPPGQ-----EGLLGFFQKAYAPFLHWT 845

DB 653 IRVEDSRADCFPCIKSARVHLANGAGPAPVYLEAPEESYLOQFIRKRVAPRLGKKT 712

QY 846 RGWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLIDYFLNRYFEVGPVYFVT 905

DB 713 KAVIMIEFLGVFAAGVALIPEVKGLDQORVAIPDDSYLIPFNLDLYEYLTGPPVYFVTR 772

QY 906 LGYNSSSAGNAICSS-AGCNNEPFTKIOVATEPPEQSYLAIPASSWVDDPIDWLTPS 964

DB 773 -EFNATDRAQQQKVCARYTTCEQMSLSNLSQERKRTVSTPTASWIDDFQWLNDPE 831

QY 965 S--CCRLYISGNKDKFCPSTVNSLCLKN---CMSITMGSVRPSVBEQPHKYLPMFLNDR 1019

DB 832 NERCC-----MERRRP-----CFANRTPAWNITLSGM-PEGDEFVYVYLKFLSAP 875

QY 1020 PNIKCPKGLAAYSTSVNLTSQGVL-ASRFMAYHKPLNSQDYTEALRAARELANITA 1078

DB 876 TNEDCPLGQASYSGLVDSRDRTIPASHFRTSHIPRSQBDFTDAYAAARRIANEISA 935

QY 1079 DLKRVPGTDAPEVPPYTTITNVFVBOYLTILPEGLFMLSCLVPTFVAVSCLLGLDLRS 1138

DB 936 -----ETGLDVFYPSVFFVFPQYASIVSLTGLGSAIGIIFIVASVLGLSLVTA 987

QY 1139 LNLLSIVMILVDTVGFNALWDISYNVSLNLNLSAVGMSVEFVSHITRSFAISTKPTWL 1198

DB 988 VVS-PTVMAIVDIIIGAMVGVSLNLSVNLNLIICVGIAEFCAHIAARAFMPSR-SCM 1045

QY 1199 ERAKE-----ATISMGSAVFAGVAMTNLPGIILVGLAKAQLQIPIFFRLNLLIT 1247

DB 1046 ERAKNRFRGRDARAWTALSNGVSGVFSGITVTKLGVFLGFTSKRKEIYFRIWVALV 1105

QY 1248 LLGLLHGLVFLPVLISYVG 1266

DB 1106 IPAATHALVFLPVALSLVG 1124

RESULT 36

Q59ZV0 CANAL

ID Q59ZV0 CANAL PRELIMINARY; PRT; 1256 AA.

AC Q59ZV0;

DT 10-MAY-2005 (Tremblrel. 30, Created)

DT 10-MAY-2005 (Tremblrel. 30, Last sequence update)

DT 10-MAY-2005 (Tremblrel. 30, Last annotation update)

DE Potential sterol homeostasis protein.

OS Name=NCR1; ORFNames=CAO19.7242;

GN Candida albicans SC5314.

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; mitosporic Saccharomycetales; Candida.

OX NCBI_TaxID=237561;

EN [1]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SC5314;

RX PubMed=15123810; DOI=10.1073/pnas.0401648101;

RA Jones T., Federspiel N.A., Chibana H., Dungan J., Kalman S.,

RA Magee B.B., Newport G., Thorstenson Y.R., Agabian N., Magee P.T.,

RA Davis R.W., Scherer S.,

RT "The diploid genome sequence of Candida albicans.";

RL Proc. Natl. Acad. Sci. U.S.A. 101:7329-7334 (2004).

RN [2]

RP NUCLEOTIDE SEQUENCE.

RC STRAIN=SC5314;

RA Dungan J., Kuo A., Newport G., Lan C.-Y., Iijima C., Adegbola O.,

RA Roberts J., Persson K., Donnelly S., Favoreto S., Tzung K.-W.,

RA Jones T., Scherer S., Agabian N.,

RT "Annotation of the Genome of Candida albicans.";

RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; AACQ0100092; BAK96004.1; -; Genomic DNA.

SQ SEQUENCE 1256 AA; 141103 MW; 89AF8C4E2ABF5286 CRC64;

Query Match

Best Local Similarity 21.2%; Score 1461.5; DB 2; Length 1256;

Matches 381; Conservative 253; Mismatches 527; Indels 181; Gaps 37;

QY 6 LRGMWLLALLRLAQSBEPTTTHQPGYCAFYDECKNPGLSGMTLSNVSLSNTPARK 65

Db	13	MRAICLLLLTINLAIAS-VSSSHKPGYCNVTGNGCKKSVFKPL-----PCAEFVPAVK	65
Qy	66	ITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNFVNLC	125
Db	66	ASQESREKLKSGCKDF-----DYICCSPEQIDILESNLKRDPLISSCPACKKNFYDFC	121
Qy	126	HNTCSNQSFLINVRVLAQAGQLPVAVVEAFYQHSFASQSDSCSRVRPAATLAV	185
Db	122	QFSCSPNESQFVEIITKTARDTGKEIVTEINQYVEPGMANQFFDCKNKFSAITN----	177
Qy	186	GTWCGVYSGALCNAORWLNFGDGTGN--GLAPLDTTEHLLPBGQAVSGGIQPLNEGVARC	243
Db	178	GYMDLIGGGAKNYSQFLKFGDEXKPLGSGSPYQINFYKLP--ETDSGLVLRNEPLDC	235
Qy	244	NESQDDVATSCQDCAASCPAIRPQALDSTFYLGQMGSLVLIILILCSFVAVVTLLV	303
Db	236	NKXE-----YKCACTDCBSCPKLPHAKDLTKKCTGVLPFCFSFIII---IWSCMIVLLG	288
Qy	304	GRVAPAR-----DKSKWVDPKKGTSLSDKLSFSFTHLLGQFFQGW-----	344
Db	289	GYHYVLAKLKKERRRSIAEDSDDESTMINPLFYAGLGKKAQKQFSSSEIGSKIQDFANI	348
Qy	345	GTWVASPLTILVLSVIVPVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHFGPF	404
Db	349	GYFCSKFPGISIGTSLAVVLLSLGLFKLQETDPVKLWSPNDPAYKNOQYFESNFGEM	408
Qy	405	FRTNQVILTAPNRSSRYRYSLLGPKNFGSILDLLELLELQELRLHLQWSEAPORN	464
Db	409	FRIEQVIVSS-----KDDGPVLNWDIVKWWFKESQLETL-----NEN	446
Qy	465	ISLQDICVAPLNDNTSLYDCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHELY	524
Db	447	VLRLSDICFKPL--DET-----CALQSFQYFQGD-----ISGLT-ETNWKSKLQS	488
Qy	525	CANAPLTPKDGITALSCMADYGAPVFPPLATGGYKGKDYSAEALIMTFSLNIPAGDP	584
Db	489	CVDSP-----VNCUPTFOQPLKNIL---FDSNDISAKAFTVTVLNSDTQEN	535
Qy	585	RLAQAQLWEAFLEEMRAPQORMAGMPQVFTFAERSLEDEINRTTAEDLPIPATYIVIP	644
Db	536	YTSNTISYEHSGKWAADLOTEVPLN-LNAYSTEISLKEELNQSNNTDITKTAISYLVWF	594
Qy	645	LYISLALGS-----YSSWSRVWDKATLGLGGVAVVLGAVMAAMGPFSLVGRSSIV	697
Db	595	IYASLALGKGLPSANLYS-----LVKTRFTLGFSSIIILLSVTASVGPFSLIGLRSTLI	649
Qy	698	ILQVWPELVLSVCAADNIFIVLEYQRLPR-RGCEPREVHIGRALGRVAPSMILCSLSEAI	756
Db	650	IAEVIFPLVLAIGDINFLIVHLELVISGNPNLALEVRISQALXKHIGSPCFISAVLQVC	709
Qy	757	CPFLGALTMPAVRTPALTSGLAVILDPLQWSAFVALLSLDSKQKQASRLDVC--CCVK	814
Db	710	MFLATSVGNPAVKNFAYTGAGAVLNFSLQMTCTFGLLALDQRLLEDNRVDVVPVWTIS	769
Qy	815	P-----OELPPQCEGLLGFQKAYAPFLHWTIRGVVLLFLALGVSLYSNCHIS	868
Db	770	PIQLQNDDEIDEPVHLEYNFSRWIGHYAPFLKTKTKPKVITLFLVWVGISLSLSPKIQ	829
Qy	869	VGLDQELALPKDSYLLDYFLNRYPEVGAPVFTVTLGYNFSSEAGMNAICSS-AGCNN	927
Db	830	LGLDQRIAPSKSYLVNYSNVVDYLVNVPVFPVVK-DLDYSESNQKICGGFSACDE	888
Qy	928	FSTQKIQYATEPEFQSYLAIPASSWDDPIDWLTG--SSCCRLYISG--PNKDKFCPT	983
Db	889	FSLANILEQBFKRSIDISMLSEPASNLWDDPFSWLNPLDQCCRFKSTVFEKTPFCSPN	948
Qy	984	VNSLNCXKCMSITMGSVRSVSQFHKYLP-----WFLND---RPNIKCPKGLAAVYSTS	1035
Db	949	APORQC-QSCYL-----NHNPPYDSSMKAPFERDFMFINDWIOEPDQPLGKAHQA	1003
Qy	1036	VNLTSDQVLASRFMAYHKPKNSQDYTEALRAARELAANITADLRKRVFGTDPAFVFPY	1095
Db	1004	ISRTTE-KIDSSYFRTSFAPLRQDQDEFINAYKSGNNIVKEITKLI-----PSMDVFAY	1055
Qy	1096	TITNVFEYQVLTLLPGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLLSTVMILAVTVGF	1155
Db	1056	SPFFIFPTQYQNVLLTVALLTVAMLIIVYVISTFLIS-SPRAASILTTITTAIMINIGV	1114
Qy	1156	MAWDDISYNAVSLINLVSAGVMSVEFVSHITRSFA-----ISTKPT	1196
Db	1115	LAWSLSLNAVTLVNLVICGVFAVEFTHLTRAYCVPKVMQFDPABEELYNLNAEPE	1174
Qy	1197	WLER-----AKEATISMGSAVFAGVAMTNLPGILVLGAKAQLOIFERL	1242
Db	1175	NTRRSSTLSLNAEPRNTKAHNALCSVGGSLISGVTLLTKLIGISVLAFTRSQIPEVYFRM	1234
Qy	1243	NLLITLLGLLHGLVFLPVILSY	1264
Db	1235	WLSLWISFVHAFVLLPVLSP	1256
RESULT 37			
Q551CS_DICDI PRELIMINARY; PRT: 1397 AA.			
AC	Q551CS;		
DT	13-SEP-2005 (TrEMBLrel. 31, Created)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last sequence update)		
DT	13-SEP-2005 (TrEMBLrel. 31, Last annotation update)		
DE	Hypothetical protein.		
GN	ORFNames=DD80202935;		
OS	Dictyostelium discoideum (slime mold).		
OC	Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.		
OX	NCBI_TaxID=44689;		
RN	NUCLEOTIDE SEQUENCE.		
RC	STRAIN=AX4;		
RA	Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,		
RA	Suegang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,		
RA	Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero P.,		
RA	Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,		
RA	Piilcher K., Chen G., Saunders D., Sodergren E., Davis P.,		
RA	Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,		
RA	Fabrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,		
RA	Cooper J., Haydock S., van Driessche N., Cronin D., Lindsay I.,		
RA	Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,		
RA	Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,		
RA	Wardrop A., Felder M., Thangavelu M., Johnson D., Knights A.,		
RA	Louisegh H., Mungall K., Oliver K., Price C., Quail M.A.,		
RA	Urushihara H., Hernandez J., Rabinowitz E., Steffen D., Sanders M.,		
RA	Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,		
RA	Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,		
RA	Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,		
RA	Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,		
RA	Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.,		
RT	"The genome of the social amoeba Dictyostelium discoideum.";		
RL	Nature 0:0-0(2005).		
CC	EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is		
CC	preliminary data.		
CC	EMBL, RAFT01000040; EAL69086.1; -; Genomic_DNA.		
KW	Hypothetical protein.		
SQ	SEQUENCE 1397 AA; 156800 MW; 3722FDB55773697F CRC64;		
Query Match 21.0%; Score 1452.5; DB 2; Length 1397;			
Best Local Similarity 26.6%; Pred. No. 1.5e-93;			
Matches 397; Conservative 247; Mismatches 465; Indels 383; Gaps 40;			
Qy	23	PYTHHQGYCAFYDECKNPELSGLMTLSNVCLSNTPARKITGDHLLILLOKICPRLY	82
Db	25	PYNTTLG---CSMYG-----VSSSFVEAKDFFPLNNT-----IAPNTCOLTH	64
Qy	83	TGPNTOACCSAKQLVSLASISITKALLTRCPACSDNFVNLCNHTCSNQSFLINVTIV	142
Db	65	PEYSIESCCNNTQTLILQTNMLVAGGIFGRCSCHVNLNLCWACSSCPYQKSFVPTKV	124

RA	Zimmer A., Zody M., Lander E.;	
RT	"The genome sequence of <i>Ustilago maydis</i> ."	
RL	Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.	
CC	-!- CAUTION: The sequence shown here is derived from an	
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is	
CC	preliminary data.	
DR	EMBL; AAC001000057; EMBL2320.1; -; Genomic_DNA.	
SK	Hypothetical protein.	
QW	SEQUENCE 1489 AA; 161971 MW; 828F4A9CA16B4BF8 CRC64;	
Query Match 20.8%; Score 1437.5; DB 2; Length 1489;		
Best Local Similarity 28.3%; Pred. No. 1.9e-92;		
Matches 428; Conservative 246; Mismatches 483; Indels 357; Gaps 51;		
QY	30 PGYAFYDECGK---NPFLSGSLMTLSNVCLNTPARKITGDHLLILQKICRLTYT-G 84	
Db	11 PGRCHKNGCKGKSFSPCLPSVDERASIPQSSFRDR-----LAATCGSEFAEG 61	
QY	85 PNTQACSAKQLVLSLEASISITKALLTRCPACSDNFVNLCHNTCSNPQSLPINVTRVAQ 144	
Db	62 P---VCCTEAQVENISANLQAEPLTSSCPACRNNFRSLFCSFTCSPPQAFVDVETQ 118	
QY	145 L-GAGOLP--AVVAYEAFQHSFARQSYDSCSRVRVPAATAATLAVGTMGVYGSALCNAQR 201	
Db	119 VTGSGKPESEAVKTYEYIDAEWKQKPFDSCKNVKFGASNGFAMDILGG--GAQTPNA-- 174	
QY	202 WLNFGDGTGNGL-APLDITFHLLRPGQAVGSGIQ-----PLNEGVAR 242	
Db	175 FLKFLGDEKPLGSPFQINF----PQSSDHRLOWAHAPSARPHNHSLAPVPFNQNRQ 230	
QY	243 CNESQDDVATSCDCAASCAPAIARPDALDFTFLGQWPGS-----L 285	
Db	231 CGDP--DLSRCTVDCDPTCTAL--PELPFS-----NQPSCACVGPWSCFTFSAILY 281	
QY	286 VLIILCSVFAVVTILLVGFVAPA-----IDLDLLELLELLELQ 448	
Db	282 MLVLVWFGVPVSAVLRGRTIALSYRTSLGSLFSQSGSFERVRMDSLSLDCVSGH 341	
QY	311 ----RDKSKVDPK-----KGTSLJ-----DKLS-----PSTHTL 336	
Db	342 PSNQRNTGSLIARGLGHYGESSASSAPDGTGRTGIGLERNDLSALGALQPRKVALNOL 401	
QY	337 LGOFFGQGTWASWPLTILVSVIPVALAGLVTELTTPVELWSAPNSQARSEKAF 396	
Db	402 LTCTFRLGLLCARHPWLTILAAVFGVGTANTGWKDFEVEDPVLRLWAPGSTAKWQDI 461	
QY	397 HQHGFQFFRTNQVILTAPNRSRYDSLLGPKNPSG-----ILDLDLLELLELQ 448	
Db	462 FDQEGFPYRPPQIFLI----DQHSYQDLSTLRNNASDPQLAALPPALSWERLLMLADLE 517	
QY	449 ERURHLQVSPQARNISLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANQT 508	
Db	518 KEVRELK--SPS---GITLQDVCLAPAG---SGTPCVQVQSTILGYFQDDP----- 558	
QY	509 LMGQTSQVD-WKHDFLYCANAPUTFGDGTALASCMADYCAPVFPFLAIGGY-KGKDYSE 566	
Db	559 -LNGGLDADSWDALCKSDNP-----VECLPSPGQPLKPNVILGGLPEGALPQ 607	
QY	567 AEALIMTFSLNYPAGDPRLAQLWEEAPL-----EEMR-----AFOR 605	
Db	608 ARAAVTVYVLLNSLNG-TTLKAAEWELELNLQEVAAAPPNKQKTSASEHHPLSVRR 666	
QY	606 RMAGMFQVTTAERLEDEINRTAEDLPFATSYIVILYISLAG----- 652	
Db	667 QQLGI-QIAFSTGVSTETGSSNTDVSIVLSTMTFYVALTYLGGSDHGVANDGS 725	
QY	653 -----SYSSWSR-----VMVDSKATLGLGGVAVV 676	
Db	726 EQGSPTEQGSYLRANVLLSYSSNGRSRLVQRTLRARSFPWNTYCVSSKFTLGLFGIV 785	
QY	677 LGAVMAAGFFSYLGRSSVLVQLVFPVLSVGADNIFIVLEYQRLPRR----- 727	
Db	786 LCSVSCAIGIFSAMGVKVLVIAEVIIPFMLAVGVNDIFLLCNEMDRQDHOHTSABFEIT 845	

QY	728 -----PG-----EPR-----EVHIGRALGRVAPSM 747	
Db	846 HSASSAPPTGVPGRSPLSPDPAVEPRGDLFMNIGASNTTGRVTVEERAARSLARVPSI 905	
QY	748 LLCSLSEACFFLGAITPMPAVRTFALTGLAVILDFLOMSFAVALLSLDSRQERASRL 807	
Db	906 LLSATTQIVAFLLGAVPMPAVNFALYAAGSMLIVAVLHCTVFIAAMALDAHVESGRI 965	
QY	808 DVCCCKV--POEL-----PPPGQEGLLGPFQKAYAPFLHLHWTITRGVVVLLFLALPG 858	
Db	966 DCMPCIKATPRQROIQLLTDVAVAKEGTLDSEIRTFAPTLERSNVKRLVVVTFGAVV 1025	
QY	859 VLSYSMCHSVGLDQELALPKDSYLLDYFLFNRYFEVGPVYFVTTLYGNYFSEAGMNA 918	
Db	1026 ISSIGVRIEMGLDQRLALPSASYLRPYDAIDVFLDVGPLYFVAT--GGEVSEKQORD 1084	
QY	919 ICSS-AGCNNFSPKIQIYATEPPEQSYLAIPASSWVDDFDWLTLP--SSCCRLYISGPN 975	
Db	1085 LCGRFTTCEPLSLASTTLEGERRRPEVSWIAEPASSWITDDFLQWLNPLDCCCKVKISDPT 1144	
QY	976 KDFECPSTVNSLCKNCM-----SITWGSVRPSVEQFHKYLPPWFLNDRPNIKCPKGG 1028	
Db	1145 --VECDPHDSPFSC-QPCFGRDPPMNITMDGL-PEGQEFYRVLKRWLESPTDOECPLGG 1200	
QY	1029 LAAYSTVNLTSD-----GOVLASRFMAYHKLNQSDYTEALRAARELANITADLRKV 1083	
Db	1201 QAYSSALSIVTDLCTGKDRVSRSHFRTYFAPLRSQSDFISALEQSQRISSESRV--- 1257	
QY	1084 PGTDPAFEPVYITNTVYFQYLTILPEGLFMLSCLVPTFPVNSCLLGLDLRSLGLNLL 1143	
Db	1258 ----GYRVFPYSLFFIFFEQYTYLLSMVQVLSAAVAIFAINTVLLG--SWRTGAVVTL 1311	
QY	1144 SIWMLVDVTGFWALWDISYNAVSLNLINLSAVGMSVEFVSHITRSF-----AIS 1192	
Db	1312 SVASAVMLVAGMGFWGQFNALTIVNLVSVCAALGVCFCAHIAAFWRAFGALPRSHPS 1371	
QY	1193 TKPTWLRARAKATISMGSAVPAGVAMTNLPGLVYLGLAKAQLIOIFFRLMLTLTLGLL 1252	
Db	1372 QKER-DERAWLALTDVGSVAVNGIFSTKLIGVGLIFTKSLDLLKLYAKTWLCLIVGGL 1430	
QY	1253 HGLVFLPVILSYVG 1266	
Db	1431 HGLILLPVLLSWLG 1444	
RESULT 39		
QY	Q9SVF0_ARATH PRELIMINARY; PRT; 1055 AA.	
ID	Q9SVF0_ARATH PRELIMINARY; PRT; 1055 AA.	
AC	Q9SVF0;	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-FEB-2005 (TrEMBLrel. 29, Last annotation update)	
DE	Hypothetical protein F22113.120 (Hypothetical protein AT4g38350).	
GN	Name=F22113.120; Synonyms=Ar4g38350;	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eurosid1; Brassicales; Brassicaceae; Arabidopsi.	
NCBI_TaxID=3702;		
ON	[1]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Bevan M., Wedler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,	
RA	Mayer K.F.X., Schueller C.;	
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
RL	[2]	
RP	NUCLEOTIDE SEQUENCE.	
RA	EU Arabidopsis sequencing project;	
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.	
RL	[3]	
RP	NUCLEOTIDE SEQUENCE.	
RA	Wedler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,	
RA	Mayer K.F.X.;	

RN [3]
RP NUCLEOTIDE SEQUENCE.
RA "Expressed genes in Ciona intestinalis";
RT Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB210578; BAE06583.1; -, mRNA.
FT NON_TER
SQ SEQUENCE 620 AA; 69500 MW; 3032C53A2A0D192 CRC64;
Query Match 20.1%; Score 1386.5; DB 2; Length 620;
Best Local Similarity 47.6%; Pred. No. 2.4e-89;
Matches 293; Conservative 98; Mismatches 187; Indels 37; Gaps 11;
QY 719 LEYQRLPRRGPPEVHHVGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSGL 778
DB 3 LEYQRLPRRGPPEVHHVGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSGL 778
QY 779 AVILDFLLQMSAFVALLSLDSKQAEASRLDVCCCKVQPELPPGQGGELLLGFFQKAYAP 838
DB 63 AIAFDLFIQITAFVLSLDARTGNGRVDVCCCKVQPELPPGQGGELLLGFFQKAYAP 120
QY 839 FLLHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGA 898
DB 121 VLMDLVRYVMVGFVGLSCWCTILCTRIITVGLDQELALPKDSYLLDYFLFNRYFEVGA 180
QY 899 PVYFTTLLGYNFSSEAGNAICSSAGCNFPSTQIKIYATEPPEQSYLAIPASSWVDDFI 958
DB 181 PVYFTTLLGYNFSSEAGNAICSSAGCNFPSTQIKIYATEPPEQSYLAIPASSWVDDFI 958
QY 959 DMLTP-----SSCRLYISGPNKDFCPSS--TVNSLN-CLKNCMSITWGSVRPSEQFH 1009
DB 241 DMLTP-----SSCRLYISGPNKDFCPSS--TVNSLN-CLKNCMSITWGSVRPSEQFH 1009
QY 1010 KYLPWFLNDRPNKCPKGLAAYSTVNL-----TSDGQ--VLASRFMAYHKLKNSQDYT 1063
DB 291 KYLPWFLNDRPNKCPKGLAAYSTVNL-----TSDGQ--VLASRFMAYHKLKNSQDYT 1063
QY 1064 EALRAARELANITADLR-----KVPCTDPAFVFPVTTINVFVEQVLTLLPGLFML 1116
DB 351 ENLIKARKLADNITLKAANKDGNLNEEDFEVFPYCLYVYVEQVLTLLPGLFML 1116
QY 1117 SLCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTFPMALWDISYNAVSLINLVSAVG 1176
DB 411 GICLIPTFAFSLILGDFYSGIITVITVIMVVDVDTAGLCSLWGVDMNAVSLINLVSAVG 1176
QY 1177 MSVEFVSHITRSPATSKPTWLERAKEATISGSAFVAGVAMTNLPGILVLGLAKAQLIQ 1236
DB 471 LSVEFISHVVRTPSLXTHISKKRVIESMTTGPVAFVAGVALLTLPGLIIVLNWATAQLIQ 530
QY 1237 IFFPRNLITLGLLGLVFLPVILSVYGVDPVNPALAEQK--RABEAAVAVMVASCP 1293
DB 531 IFFPRNLITLGLLGLVFLPVILSVYGVDPVNPALAEQK--RABEAAVAVMVASCP 1293
QY 1294 NHPSRVSTADNIYV 1308
DB 591 QEKQCKTKNKNVYN 605

RESULT 41
Q618V4_CABBR
ID Q618V4_CABBR PRELIMINARY; PRT; 1382 AA.
AC Q618V4_CABBR
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14473.
GN Name=CBG14473;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.

RG The C. briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CRAC01000068; CAB68601.1; -, Genomic DNA.
DR GO; GO:0016021; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 1382 AA; 155281 MW; CA269DAB5C9CE527 CRC64;

Query Match 19.7%; Score 1359; DB 2; Length 1382;
Best Local Similarity 27.8%; Pred. No. 6.3e-87;
Matches 357; Conservative 253; Mismatches 524; Indels 148; Gaps 36;
QY 74 LQKICPLRYTGPNTQACCSAKQLVSLSEASLITKALLTRCPACSDNFVNLHCHNTCSNQ 133
DB 61 MMEFCPHMGGDN-KUCCTPSQAEGLTKQIAQARHILGRCPSCFDNPAKLWCEFTCSNQ 119
QY 134 SLFINVTRVAQLAGAGQLPVAVYEAIFYQHS-----PAPQSYDSCSRV-- 177
DB 120 QDFVSIEMKPIEKKE-----GFSEYQPAEAVYVTVYRLSTEFABGMPASCKDVTFGG 174
QY 178 -PAAATLAVGTMCGVSGALCNQRLNFOGDTGNGLA-PLDITFHLLEPQAVGSGIQP 235
DB 175 QPALRVMCISIPC-----TLTN---WLDFTGTQNLDSLIPINTNFILYDPAKTPQSDRSI 226
QY 236 -LNEGVARCNESQDDVATCSCDCAAPALARPQALDSTFFYLGOM-----PGSLVLI 289
DB 227 FMNVFTGDSQARPGHPACKSEC--NKEEYAKMNLDDGKTSQCNVHGIACLNIFV 284
QY 290 ILCVFAVATLLVGFVRPAPARKSKMVDPKKGTSLSDKLSFSTHTLLGQFQGWGTWA 349
DB 285 MLAGISLAVLLCVGFVFTSYDEBYTVLRQNAQGESPKRNRIRT-----GAWIH 335
QY 350 SW-----PLTLVLSVIPVVALAAGLVTELTTPDELWSAPNSQARSEK 394
DB 336 NFMENNARDIGMAGRNPKSHFFIGCAVLIFCLPGMIYHOESTNVDMWSSPGRARQEE 395
QY 395 AFHQHGFPPRTNQVILTAPNRSSRYSDLSLLGPKNPSGILDLLLELELQERLRL 454
DB 396 TVFNANFORPORYQIIMLLS-NR-EFQNGKLYGP-----VFHKDIFEEFLDILNAIKNI 448
QY 455 QVMSPEAQRNLSLODICYAPLNPDNTSLYDCINSLLOVFO-NRNTLLLLTANQTLMGQT 513
DB 449 STKDAEG-RTVTLDDVCYRPMGP-----YDCLIMSPTNYFOARGKENLELKEEIVSDED 503
QY 514 SQVD-----WKDHLVCANAPLTFKDGTLALASCMADYGPAPVFPFLAIGYKGD 563
DB 504 DAFDYFSSEGTDEMMNHIAACIDQPMQOK--TKSGLSCFGTYGGSAPNM-VFGRNTYN 560
QY 564 YSEAEALIMTFLNNYPAGDPRLAQALKEBAFLFEEMAFORRMAGNFQVTF--AERSL 621
DB 561 YQANSMVMTFLVTO--RTEPEIQAELEWEKEFIKCKDYREKSP--KVTFSPFAERSI 615
QY 622 EDEINRTTAEDLPFIPATSYVIFLYISALASY-----SSWSRVWVDSKATILGCVAVV 676
DB 616 PDEIEKDAKDEIVTVVIALAFILGYVTFSLGRYFVCENQLWS-ILVHSRICLGMLSVIN 674
QY 677 LGAVNAAGFFSYLGIKRSVLQVVPFLVSVGADNIFIFVLEYORL----PRRPGEP 732
DB 675 LLSFSCSGWIFSMFGHPVKNALVQFVTVLUGVCRFMVKVYTAQLAVSPNYSPODC 734
QY 733 EVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTALTSGLAVILDFLLQMSAPV 792
DB 735 PAIVGVWAGTWPAMFSSSLGCAFSFFIGGFTDLPAIRTFCLYAGLAVILDFVVLHCTIPL 794

Db 191 -MC---TSPCTLTNLEFIGNLNDLNIPHTHFKLLYDPIKTPSPDRSTMTNVTGCD 236
Qy 245 ESQGDVATSCQDCAACPAIARQALDSTFYLGQM-----PGSLVLIILCSVFAVVT 299
Db 237 KSARVGFACSTSEC--NKBEYANLIDDDGKTSQTCNVHGIACLAIFVMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARKSKVDPKGTSLSDKLSFSTHTLLGQFF 341
Db 295 LLCVGFVFTSYDEDTYTNLRQTSGESPKNRK-----RTGAWI-----HNFMENNA 342
Qy 342 CGMGTWASWPLTILVLSVIPVALAAGLVFTLTDVPVLSWSPNSQARSEKAFDHOHF 401
Db 343 RDIGMAGNPKSHFFIGCAVLIFCLPGWYHKESTNVVDMWSPSRARQEMVFNANF 402
Qy 402 GPFFRTNQVILTPAPNRSRYRDSLLGPKNFGILDLDLLELELQRLRHLQWSPEA 461
Db 403 GRPQRYQQIMLL--SHRDFQSGKLYG-----VPHKDFBELFDILNAIKNISQDSG 455
Qy 462 QBNISLDICVAPLNPDNTSLYDCCINSLLQYFQNNRTLLLTANQT-----L 509
Db 456 -RTITLDDVCRPMGFG---YDCLIMSPNTYFQNKHELDKMSKETVSEDDDDAFYF 510
Qy 510 MGQTSQVQDKHFLYCANAPLTFKDGTTALALSCMADYGAPVFPFLAIGYKRGQSEAE 569
Db 511 SSEATDEWNNHMAACIDOPMSQK--TKSGLSGMYGPGSAPNM-VFGKSNTHQAANS 567
Qy 570 LIMTSLNNYPAGDRLQAQKWEBAFLIEMRAFORRMAGMPQVFTTFAERSLEDBINTT 629
Db 568 IMWTILVQ--RTEPEIQAKELWEKFLKFCKEYREKSPKVI-FSFMERSITDIENDA 624
Qy 630 AEDLPIFATSYIVILYISLAGSY-----SSSRVMDVSKATLGLGVAVVGLVWMAAM 684
Db 625 KDEIVTWIALAFLIGYVTFSGRYFVCENQLWS--ILVHSRCLGMLSVIINLLSFCSW 683
Qy 685 GFPSYLGIRSSLVILQVFPVLVSAGADNIFTVLEY--QRUP---RRPGPREVHIGRA 739
Db 684 GIFSMFGIHPVKNALVQVFFVTLIGVCTFMVKYACQVRVMPMSDPQCEI-VGMV 742
Qy 740 LGRVAPMLLCSLSEACFFLGAALTMPAIRTALTSLGLAVLDFLLOMSAFVALLSDS 799
Db 743 MAGTMPAMFSSILGCASFPGFIDLPARTFCLYAGLAVLIDVVLHCTIFLFWMDT 802
Qy 800 KQEQASRLDVCCVQKQELPPQCGEGL-----LLG-----FFQKAVAPFLHWI 844
Db 803 QRELNG-----KP-EFFFPYQIKDLGAVLIGQRATDTFTQTFHQVAPFLMRM 853
Qy 845 TRGVULLFLALFGVSLYSWCHISVGLDQELALPKDSYLLDFLFLNRYFEYGAPVYFT 904
Db 854 TRITGIIIFIASPITTVILSSKISVGFQDSMAFTEKSYISTHRYLDKFDVGPVPVFTV 913
Qy 905 TLGYNFSBAGNAICSSAGCNPSPSTQIKVATEPPEQSYLAIPASSWDDFDWLT-P 963
Db 914 DGELDWHRPDVQNKFTFPFGCSDTSFGNTMNVAVGHTQYLSGEMYNWIDYLEWISRK 973
Qy 964 SSCRLYISGPNKDKCPSTVNSL-----NCLNKM-----SITWGSV---RPSVEOPH 1009
Db 974 SPCCVKYVHDEN--TFCSTNRNKSALDDKACRTCMDPDFVANSYPKSITWYRPSLEVY 1031
Qy 1010 KYLPWFLNDRPNIKCPKGGIAAYSTSVNLTSDQVLASRFMAYHKPL--KNSQDYTEALR 1067
Db 1032 RHLRHLEDTPNSECVGGRASPKDAISFTSRGICQSQFMTHFKKLSISNSDDFIKAM 1091
Qy 1068 AARELANITADLRKVPGTDPAPVEFPYITNVFYEQYITLPEGLFMLSCLCLVPTFAVS 1127
Db 1092 TARMVSRRLERSI-----DDTAHVAFAYSKIFPEYEQYSTIMPLTQTFITVWVGFI 1145
Qy 1128 CLLGLDLRGLNLLSIYMLIVDTVGEWALWDISYNAVSLNLVSAVGMSVEFVSHTR 1187
Db 1146 CVTLGIDVKGACAVICQVSNYTHIVAFMTIFIPVNLASATNLVMSGILIEFSVNVLK 1205
Qy 1188 SFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLT 1247

Db 1206 GYACSLRQRAKRAESTVSGSIGPIILSGPVVTVMAGSTMFLSGAHLQIITVYFFKFLITI 1265
Qy 1248 LLGLLHCLVLPVLISVGPDPVNPALALEQKRAEEAAVAVMVASCNHPRSVSTADNIYV 1307
Db 1266 VSSAVHALIITLPIILAPGSGRSGHSGSTSTNDNDQHDACVLS--PTAESHISNVEGIL 1323
Qy 1308 N-----HSFEGSIKAGAI 1321
Db 1324 NRPSLLDASHILDPLLKAEGL 1345
RESULT 43
Q12200 YEAST
ID Q12200 YEAST PRELIMINARY; PRT; 1170 AA.
AC Q12200;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)
DT 10-MAY-2005 (TrEMBLrel. 30, last annotation update)
DE Hypothetical protein YPL006W (Ypl006wp).
GN Name=NCRI; OrderedLocusNames=YPL006W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=AB972;
RA Badcock K., Bowman S., Churcher C.M., Pearson D., Rajandream M.A.,
RA Walsh S.V., Barrrell B.G.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=89057455; PubMed=3143101;
RA O'Hara P.J., Horowitz H., Eichinger G., Young E.T.;
RT "The yeast ADK6 gene encodes homopolymeric amino acid sequences and a
RT potential metal-binding domain";
RL Nucleic Acids Res. 16:10153-10169(1988).
RN [3]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=91092258; PubMed=2265610;
RA Gerrig S.L., Spencer F., Hieter P.;
RT "The CHL 1 (CTF 1) gene product of Saccharomyces cerevisiae is
RT important for chromosome transmission and normal cell cycle
RT progression in G2/M";
RL EMBO J. 9:4347-4358(1990).
RN [4]
RP NUCLEOTIDE SEQUENCE.
EX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansoorge W.,
RA Busey H., Aparicio A., Barrrell B.G., Badcock K., Benes V.,
RA Botstein D., Bowman S., Brueckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Dietrich F.S., Delius H., DiPaolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Hunnicke-Smith S., Hyman R.W., Johnston M., Kalman S., Kleine K.,
RA Komp C., Kurdi O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mirtipati S., Moestl D.,
RA Mueller-Auer S., Namath A., Nentsch U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.M., Purnelle B., Rajandream M.A., Reckmann S.,
RA Rieger M., Riles L., Roberts D., Schaefer M., Scharte M., Scherens B.,
RA Schramm S., Schroeder M., Sdicu A.-M., Tettelin H., Urrestazu L.A.,
RA Ushinsky S., Vierendeels F., Visiers S., Voss H., Walsh S.V.,
RA Wambutt R., Wang Y., Wedler E., Wedler H., Winnett E., Zhong W.-W.,
RA Zollner A., Vo D.H., Hani J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI";
RL Nature 387:103-105(1997).
RN [5]
RP NUCLEOTIDE SEQUENCE.
RA Jia Y., Cherry J.M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [6]
RP NUCLEOTIDE SEQUENCE.


```
DR Pfam: PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40801 MW; EA1D521AD61B9975 CRC64;

Query Match 12.7%; Score 876; DB 2; Length 371;
Best Local Similarity 47.2%; Pred. No. 1.9e-53;
Matches 180; Conservative 70; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGNNAICSSAGCANNFSFTQKIYATEPPEQSYLAIPASS 952
Db 1 YLSVGAPVYFVVKDQGNVYTDAGANQICGGMGNNNSLIEQIARMSKMPNYSHIAYPASS 60

QY 953 WDDDFIDWLTP--SSCCRLYISGPNKDKFCPSTVNSLNCNKMCI--TWGSRVRSVEQPH 1009
Db 61 WLDDYDFWLKPOSSCCRHDNTG--EEDVFCNATVWTSTCIA--CRSAQESANQSRPTDFEFM 118

QY 1010 KYLPWFNDLRPNKICPKGGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYTEAL 1066
Db 119 KFLPWFNDLRPNKICPKGGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYTEAL 1066

QY 1067 RAARELANITADLRKVPETDPAFVFPYTTITNVFEQYLTILPEGLFMLSLCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAEVFPYFVFPYFVEQYLTIVHDTIENLGVSLAAIFV 230

QY 1127 SCLLGLDLRSGLNLLSVMLVDTVGFMAWDISYNAVSLNLVSAVGMVSVEFVSHIT 1186
Db 231 VFLLGFDLLSAVIVVLTLLDLMFGAMYLWNIPLNAVSLNLVMAVGLSVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGS--AVFAGVAMTNLPGLVLGLAKAQLIQIFPRLNL 1244
Db 291 RAPALSQRITRVARAEALAEIGSSVSVLSGITLTKFVGIVILAFSKQIFKVFYFRMYL 350

QY 1245 LITLLGLHLGLVFLPVILSYV 1265
Db 351 CVVVLGAGHGLVFLPVILSYI 371

RESULT 51
Q66NR8 CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Japan 1;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667367; AAU04386.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON TER 371 371
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40801 MW; 5BDE9D3DDF28CE00 CRC64;

Query Match 12.6%; Score 873; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGNNAICSSAGCANNFSFTQKIYATEPPEQSYLAIPASS 952
Db 1 YLSVGAPVYFVVKDQGNVYTDAGANQICGGMGNNNSLIEQIARMSKMPNYSHIAYPASS 60

QY 953 WDDDFIDWLTP--SSCCRLYISGPNKDKFCPSTVNSLNCNKMCI--TWGSRVRSVEQPH 1009
Db 61 WLDDYDFWLKPOSSCCRHDNTG--EEDVFCNATVWTSTCIA--CRSAQESANQSRPTDFEFM 118

QY 1010 KYLPWFNDLRPNKICPKGGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYTEAL 1066
Db 119 KFLPWFNDLRPNKICPKGGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYTEAL 1066

QY 1067 RAARELANITADLRKVPETDPAFVFPYTTITNVFEQYLTILPEGLFMLSLCLVPTFAV 1126
Db 179 RSANEIAEQIS-----QNTTAEVFPYFVFPYFVEQYLTIVHDTIENLGVSLAAIFV 230

QY 1127 SCLLGLDLRSGLNLLSVMLVDTVGFMAWDISYNAVSLNLVSAVGMVSVEFVSHIT 1186
Db 231 VFLLGFDLLSAVIVVLTLLDLMFGAMYLWNIPLNAVSLNLVMAVGLSVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGS--AVFAGVAMTNLPGLVLGLAKAQLIQIFPRLNL 1244
Db 291 RAPALSQRITRVARAEALAEIGSSVSVLSGITLTKFVGIVILAFSKQIFKVFYFRMYL 350

QY 1245 LITLLGLHLGLVFLPVILSYV 1265
Db 351 CVVVLGAGHGLVFLPVILSYI 371

RESULT 51
Q66NR8 CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 14;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667361; AAU04380.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrflvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON TER 371 371
FT NON TER 371 371
SQ SEQUENCE 371 AA; 40815 MW; CD137A42310A413C CRC64;

Query Match 12.7%; Score 874; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 2.7e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGNNAICSSAGCANNFSFTQKIYATEPPEQSYLAIPASS 952
```

```
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLITLPEGLFMLSCLVPTFAV 1126
DB 179 RSANEIAEQIS-----QNTTAEVPFYSVFPYEQYLTIVHTDITFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRSLGMLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGVMSVFPVSHIT 1186
DB 231 VFPLGFDLLSAVIVVVTLLDLMFGAMYLWNIPLNAVSLVNLVMAVGSVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNL 1244
DB 291 RAPALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAPSKSIQIFKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
DB 351 CVMVLGAGHGLVFLPVLSYI 371

RESULT 53
Q66NR4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR4;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 18;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667365; AAU04384.1; -; Genomic_DNA.
DR EMBL; AY667365; AAU04385.1; -; Genomic_DNA.
DR EMBL; AY667352; AAU04371.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40786 MW; E81FCD07C34BF83 CRC64;

Query Match 12.6%; Score 873; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YPEVGAPVYFVTLGYNFSSAGMNAICSSAGCANNFSPQKIQYATEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVKDQNYTDAAGANQICGGMGCNNLSLIEQIARMSKMPNYSHIAYPASS 60
QY 953 WDDFDLWLP--SSCCRLYISGPNKDKFCPSTVNSLNCNKCMSI--TWGSRVPSVEQFH 1009
DB 61 WDDYDFDLWKQSSCCRDHNTG--KEDVFCNAIVTSTSCIA--CRSAQESANQSRPTPDEFM 118
QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYREAL 1066
DB 119 KFLPWLNDNPETKCAKGHAAYGTSVKVIDEGKSRVGTGTFMAYHTLTSTKDFIGCL 178
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLITLPEGLFMLSCLVPTFAV 1126
DB 179 RSANEIAEQIS-----QNTTAEVPFYSVFPYEQYLTIVHTDITFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRSLGMLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGVMSVFPVSHIT 1186
DB 231 VFPLGFDLLSAVIVVVTLLDLMFGAMYLWNIPLNAVSLVNLVMAVGSVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNL 1244
DB 179 RSANEIAEQIS-----QNTTAEVPFYSVFPYEQYLTIVHTDITFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRSLGMLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGVMSVFPVSHIT 1186
DB 231 VFPLGFDLLSAVIVVVTLLDLMFGAMYLWNIPLNAVSLVNLVMAVGSVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNL 1244
DB 291 RAPALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAPSKSIQIFKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
DB 351 CVMVLGAGHGLVFLPVLSYI 371

RESULT 55
Q66NR7_CIOIN PRELIMINARY; PRT; 371 AA.
ID Q66NR7_CIOIN
```

```
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLITLPEGLFMLSCLVPTFAV 1126
DB 179 RSANEIAEQIS-----QNTTAEVPFYSVFPYEQYLTIVHTDITFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRSLGMLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGVMSVFPVSHIT 1186
DB 231 VFPLGFDLLSAVIVVVTLLDLMFGAMYLWNIPLNAVSLVNLVMAVGSVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNL 1244
DB 291 RAPALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAPSKSIQIFKVFYFRMYL 350
QY 1245 LITLGLHLGLVFLPVLSYV 1265
DB 351 CVMVLGAGHGLVFLPVLSYI 371

RESULT 53
Q66NR4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q66NR4;
DT 25-OCT-2004 (TremBLrel. 28, Created)
DT 25-OCT-2004 (TremBLrel. 28, Last sequence update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 18;
RX PubMed=15545499; DOI=10.1101/gr.3199704;
RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667365; AAU04384.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40786 MW; E81FCD07C34BF83 CRC64;

Query Match 12.6%; Score 873; DB 2; Length 371;
Best Local Similarity 47.0%; Pred. No. 3.2e-53;
Matches 179; Conservative 71; Mismatches 113; Indels 18; Gaps 7;

QY 893 YPEVGAPVYFVTLGYNFSSAGMNAICSSAGCANNFSPQKIQYATEPPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVKDQNYTDAAGANQICGGMGCNNLSLIEQIARMSKMPNYSHIAYPASS 60
QY 953 WDDFDLWLP--SSCCRLYISGPNKDKFCPSTVNSLNCNKCMSI--TWGSRVPSVEQFH 1009
DB 61 WDDYDFDLWKQSSCCRDHNTG--KEDVFCNAIVTSTSCIA--CRSAQESANQSRPTPDEFM 118
QY 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDG---QVLASRFMAYHKPLKNSQDYREAL 1066
DB 119 KFLPWLNDNPETKCAKGHAAYGTSVKVIDEGKSRVGTGTFMAYHTLTSTKDFIGCL 178
QY 1067 RAARELAANITADLRKVPCTDPAPFVPPYITINVFYEQYLITLPEGLFMLSCLVPTFAV 1126
DB 179 RSANEIAEQIS-----QNTTAEVPFYSVFPYEQYLTIVHTDITFNLGVSLLAAIFV 230
QY 1127 SCLLGLDLRSLGMLLSIVMLIVDTVGFMALWDISYNVSLINLVSAGVMSVFPVSHIT 1186
DB 231 VFPLGFDLLSAVIVVVTLLDLMFGAMYLWNIPLNAVSLVNLVMAVGSVFECAHIT 290
QY 1187 RSPAISTKPTWLERAKEATISMGs--AVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNL 1244
DB 291 RAPALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAPSKSIQIFKVFYFRMYL 350
```

```
AC Q6NR7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 15, West coast 16, Japan 2, Japan 3, Japan 4, and
RC West coast 6;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RX Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RA "Intraspecies sequence comparisons for annotating genomes.";
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667362; AAU04381.1; -; Genomic_DNA.
DR EMBL; AY667363; AAU04382.1; -; Genomic_DNA.
DR EMBL; AY667368; AAU04387.1; -; Genomic_DNA.
DR EMBL; AY667369; AAU04388.1; -; Genomic_DNA.
DR EMBL; AY667353; AAU04372.1; -; Genomic_DNA.
DR EMBL; AY667370; AAU04389.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40787 MW; D05762F36018549 CRC64;

Query Match 12.6%; Score 869; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 6.1e-53;
Matches 178; Conservative 72; Mismatches 113; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVVKDQNYTDAAGANQICGGMGNNSLIEQIARMSKPNYSHIAYPASS 60

QY 953 WDDFDIDLTP--SSCRLYISGNPKDKFCPSTVNSLNCNKMSI--TMGSVRPSVEQFH 1009
DB 61 WDDYDFDLKPKQSSCCRHDNTG-EEDVFCNATVVTSTCIA-CRSAQESANQSRTPDEFM 118

QY 1010 KYLPWFLNDRPNKCPKGLAAYSTSVNLTSQG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNLNDNPETKCAKGHAAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKTSDFIGCL 178

QY 1067 RAARELAANITADLRKVGTDPAFVFPYTTITNVFEQYLTITLPEGLFMLSCLVPTPAV 1126
DB 179 RSANKIAEIS-----QNTTAEVFPYSVFYFYEQYLTIVHDTIFNLGVSLLAIFVV 230

QY 1127 SCLLGLDLRSLGLNLISVMILVDTVGFPMALWDISYNAVSLINLVSAGMSVEFVSHIT 1186
DB 231 VFLLGFDDLSSAVIVVVTILLDDMGAMYLWNIPLNAVSLVNLVMAVGISVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGS--AVFAGVAMTNLPGILVLGLAKAQLIQIFPRML 1244
DB 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFSKSIQFKVYFPRMYL 350

QY 1245 LITLLGLLHGLVFLPVLISV 1265
DB 351 CVVVLGAGHGLVFLPVLISYI 371

RESULT 56
Q6NS4_CIOIN
ID Q6NS4_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q6NS4;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
```

```
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 8;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RX Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RA "Intraspecies sequence comparisons for annotating genomes.";
RT "Intraspecies sequence comparisons for annotating genomes.";
RL Genome Res. 14:2406-2411(2004).
DR EMBL; AY667355; AAU04374.1; -; Genomic_DNA.
DR InterPro; IPR001036; Acrlflvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
FT NON_TER 1
FT NON_TER 371
SQ SEQUENCE 371 AA; 40821 MW; EA1D4F58DFB98D95 CRC64;

Query Match 12.6%; Score 869; DB 2; Length 371;
Best Local Similarity 46.7%; Pred. No. 6.1e-53;
Matches 178; Conservative 71; Mismatches 114; Indels 18; Gaps 7;

QY 893 YFEVGAPVYFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPEQSYLAIPASS 952
DB 1 YLSVGAPVYFVVKDQNYTDAAGANQICGGMGNNSLIEQIARMSKPNYSHIAYPASS 60

QY 953 WDDFDIDLTP--SSCRLYISGNPKDKFCPSTVNSLNCNKMSI--TMGSVRPSVEQFH 1009
DB 61 WDDYDFDLKPKQSSCCRHDNTG-EEDVFCNATVVTSTCIA-CRSAQESANQSRTPDEFM 118

QY 1010 KYLPWFLNDRPNKCPKGLAAYSTSVNLTSQG---QVLASRFMAYHKPLKNSQDYTEAL 1066
DB 119 KFLPWFNLNDNPETKCAKGHAAAYGTSVKVIDEGKSRVGATSFMAHYHTLTKTSDFIGCL 178

QY 1067 RAARELAANITADLRKVGTDPAFVFPYTTITNVFEQYLTITLPEGLFMLSCLVPTPAV 1126
DB 179 RSANEIAEQIS-----QNTTAEVFPYSVFYFYEQYLTIVHDTIFNLGVSLLAIFVV 230

QY 1127 SCLLGLDLRSLGLNLISVMILVDTVGFPMALWDISYNAVSLINLVSAGMSVEFVSHIT 1186
DB 231 VFLLGFDDLSSAVIVVVTILLDDMGAMYLWNIPLNAVSLVNLVMAVGISVEFCAHIT 290

QY 1187 RSPAISTKPTWLERAKEATISMGS--AVFAGVAMTNLPGILVLGLAKAQLIQIFPRML 1244
DB 291 RAFALSQRITRVARAEALAEIGSSVLSGTLTKFVGIVILAFSKSIQFKVYFPRMYL 350

QY 1245 LITLLGLLHGLVFLPVLISV 1265
DB 351 CVVVLGAGHGLVFLPVLISYI 371

RESULT 57
Q6NT1_CIOIN
ID Q6NT1_CIOIN PRELIMINARY; PRT; 371 AA.
AC Q6NT1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Patched (Fragment).
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=West coast 1;
RC PubMed=15545499; DOI=10.1101/gr.3199704;
RX Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
RA Keys D.N., Rubin E.M.;
RA "Intraspecies sequence comparisons for annotating genomes.";
RT "Intraspecies sequence comparisons for annotating genomes.";
```


Genome Res. 14:2406-2411(2004).
 EMBL; AY667348; AAU04367.1; -; Genomic_DNA.
 InterPro; IPR001036; Acrlflvin_res.
 InterPro; IPR003392; Patched.
 Pfam; PF02460; Patched; 1.
 PRINTS; PR00702; ACRIPLAVINRP.
 FT NON_TER 371 371
 SQ SEQUENCE 371 AA; 40863 MW; B3745263AGAB8B5 CRC64;
 Query Match 12.5%; Score 867; DB 2; Length 371;
 Best Local Similarity 46.7%; Pred. No. 8.4e-53;
 Matches 178; Conservative 71; Mismatches 114; Indels 18; Gaps 7;
 QY 893 YFEGAPVYFVTTLGYNFSEAGMAICSSAGCNFFSFTQKIYATPEPPEQSYLAIPASS 952
 Db 1 YLSVGAPVYFVKDQGNQYTDAGANQICGGMGNNLSLIEQIARMSKMNYSIAIYAPASS 60
 QY 953 WDDFDLWTFP-SSCRLYISGNKDKPCFSTVNSLNCNCKMSI--TWGSVRPSVEQPH 1009
 Db 61 WLDDYFDWLKPOSSCCRHDNTG-BEDVFCNATVWSTSCIA-CRSQAESANQSRPTDPEFM 118
 QY 1010 KYLPWFLNDRPNKCPKGLAAYSTSVNLTSDG---QVLASRFMAVHKPLKNSQDYTEAL 1066
 Db 119 KFLPFLNDRPNKCPKGLAAYSTSVNLTSDG---QVLASRFMAVHKPLKNSQDYTEAL 178
 QY 1067 RAARELAANITADLRKVPCTDPAFEVFPYITITNVFYEQYLTILPEGLFMLSCLVPTFAV 1126
 Db 179 RSANEIAEQIS-----QNTTAEVFPYVSVFVFEQYLTIVHDTIFNLGSLAAIFV 230
 QY 1127 SCLLGLDRLSGLLNLLSVMILVDTVGFMAWLDISYNAVSLINLSVAGMSVFPVSHIT 1186
 Db 231 VFFLLGDFPLSAIVVVTLLILDMFGAWLWNLPLNAVSLVNLVMAVGISVFCAHIT 290
 QY 1187 RSPALSTKPTWLERAKEATISNGS--AVFAGVAMTNLPGILVLGLAKAQLIOIFPRNL 1244
 Db 291 RAPALSORITVARAEALAEIAGSVSVLSGITLTKFVGIVILAFSKSIQPKVFYFRMYL 350
 QY 1245 LITLLGLHGLVPLVILSYV 1265
 Db 351 CVVVLGAGHGLVPLVILSYI 371

RESULT 58

Q9C8F0 ARATH
 ID Q9C8F0_ARATH PRELIMINARY; PRT; 524 AA.
 AC Q9C8F0_
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein T8D8.3 (Fragment).
 GN Name=T8D8.3;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_TaxID=3702;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
 RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RA Town C.D., Kaul S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC025815; AAG51318.1; -; Genomic_DNA.
 DR GO; GO:0016020; C-membrane; IEA.
 DR GO; GO:0008158; F.hedgehog receptor activity; IEA.
 DR InterPro; IPR003392; Patched.
 DR Pfam; PF02460; Patched; 1.
 KW Hypothetical protein.

FT NON_TER 1 1
 SQ SEQUENCE 524 AA; 58105 MW; 4414D9C5D4FD0333 CRC64;
 Query Match 12.4%; Score 858.5; DB 2; Length 524;
 Best Local Similarity 38.1%; Pred. No. 5.3e-52;
 Matches 188; Conservative 94; Mismatches 152; Indels 59; Gaps 13;
 QY 822 GQGE---GLLGGFFQKAYAPFLHMTTRGVVLLFP--LALFGVSLYSMCHISVGLDQELA 876
 Db 24 GVGQRKAGLLTRYMKVHAPVLSHWIVKIVVIAFPFGLAWAGIALST--RIEPLSQIV 81
 QY 877 LPKDSYLLDYFLNRYFEVGAIPVFTVTLGNFNSSEA-GMNAICSSAGCNFFSFTQKIY 935
 Db 82 LPQDSYLOGYFNINSTRYLRIIGPPLFYVLK-NYNYSSESRTNQLCSINKCNPNLSLNEIA 140
 QY 936 YATEPEQSYLAIPASSWDDFDLWLTSS--CCRLYISG---PNKDK--PCPSTVNSLN 988
 Db 141 RASLTFELSYIAKPAASMLDDFLVWLSPEAFCCCRKFTNGTCPPDDQPPCCPPGQTS 200
 QY 989 CLKNCMSIT-----MGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLT-- 1039
 Db 201 LSEVKDCCTTCFRHADLSSDRPSTTQPKKLPWFLNALPSADCAKGGHAYSSVDLQY 260
 QY 1040 SDQVLAERFMAHYHPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITN 1099
 Db 261 ANGIQASSFRYTHPLNKQVDFVNSMRAAQBSAKVSRSLK-----MBIYPYSVFY 312
 QY 1100 VPEYOYLTPGLPMLSLCIVPTFAVSCLLGLDRLSGLLNLLSVMILVDTVGFMAW 1159
 Db 313 MFFEYGLDITKWTALNLSIAIAVFW-CLITCSFWSSAIIILLVIAMIIIDLGVMAVF 371
 QY 1160 DISYNAVSLINLSVAGMSVFPVSHITRSFALSTKPTWL----- 1198
 Db 372 HQLNALSVNLSVGLAVECVHITHAFSVSLSVCTSIHKKQFAIVTLTVRNAQIST 431
 QY 1199 ----BRAKEATISMSGSAFVAGVAMTNLPGILVLGLAKAQLIOIFPRNLTLTGLH 1254
 Db 432 GDRNRMKEALGGMGASVFGITLKVGVILGFSRSEVFVYVYFKYMLALVLLGLFLHG 491
 QY 1255 LVFLPVLVSYVCP 1267
 Db 492 LVFLPVLVSMFGP 504

RESULT 59

Q66NS0 CIOIN
 ID Q66NS0_CIOIN PRELIMINARY; PRT; 371 AA.
 AC Q66NS0;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Patched (Fragment).
 OS Ciona intestinalis.
 OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
 OC Phlebobranchia; Cloniadae; Cliona.
 OX NCBI_TaxID=7719;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC STRAIN=West coast 12;
 RX PubMed:15545499; DOI=10.1101/gr.3199704;
 RA Boffelli D., Weer C.V., Weng L., Lewis K.D., Shoukry M.I., Pachter L.,
 RA Keys D.N., Rubin E.M.;
 RT "Intraspecies sequence comparisons for annotating genomes."
 RL Genome Res. 14:2406-2411(2004).
 DR EMBL; AY667359; AAU04378.1; -; Genomic_DNA.
 DR InterPro; IPR003392; Patched.
 DR Pfam; PF02460; Patched; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 371 AA; 40633 MW; 168F3A36B45EEA3 CRC64;
 Query Match 12.4%; Score 854; DB 2; Length 371;
 Best Local Similarity 46.5%; Pred. No. 7e-52;

RA McLay K.E., McMurray A., Milne S., Nickerson T., Nisbett J.,
RA Nordstiek G., Pearce A.V., Peck A.I., Porter K.M., Pandian R.,
RA Pelan S., Phillimore B., Povey S., Ramsey Y., Rand V., Scharfe M.,
RA Sehra H.K., Showkeen R., Sims S.K., Skuce C.D., Smith M.,
RA Steward C.A., Swarbrick D., Syncamore N., Tester J., Thorpe A.,
RA Tracey A., Tromans A., Thomas D.W., Wall M., Wallis J.M., West A.P.,
RA Whitehead S.L., Willey D.L., Williams S.A., Wilming L., Wray P.W.,
RA Young L., Ashurst J.L., Coulson A., Blocker H., Durbin R.,
RA Sulston J.B., Hubbard T., Jackson M.J., Bentley D.R., Beck S.,
RA Rogers J., Dunham I.,
RT "DNA sequence and analysis of human chromosome 9.";
RL Nature 429:369-374(2004).
RN [3]
RN NUCLEOTIDE SEQUENCE [MRNA] OF 152-1447.
RC TISSUE=Fetal brain;
RX MEDLINE=96218118; PubMed=8647801; DOI=10.1074/jbc.271.21.12125;
RA Hahn H., Christiansen J., Wicking C., Zaphiropoulos P.G.,
RA Chidambaram A., Gerrard B., Vorechovsky I., Bale A.E., Toftgard R.,
RA Dean M., Wainwright B.J.;
RT "A mammalian patched homolog is expressed in target tissues of sonic
RT hedgehog and maps to a region associated with developmental
RT abnormalities.";
RL J. Biol. Chem. 271:12125-12128(1996).
RN [4]
RN VARIANTS BCNS ARG-509; VAL-509; GLN-816 DEL AND TYR-1132.
RX MEDLINE=96438609; PubMed=8940969;
RA Chidambaram A., Goldstein A.M., Gailani M.R., Gerrard B., Bale S.J.,
RA DiGiovanna J.J., Bale A.E., Dean M.;
RT "Mutations in the human homologue of the Drosophila patched gene in
RT Caucasian and African-American nevoid basal cell carcinoma syndrome
RT patients.";
RL Cancer Res. 56:4599-4601(1996).
RN [5]
RN VARIANTS BCNS TYR-513 AND ARG-1069.
RX MEDLINE=97136566; PubMed=8981943;
RA Wicking C., Shanley S., Smyth I., Gillies S., Negus K., Graham S.,
RA Suthers G., Hailes N., Edwards M., Wainwright B.J.,
RA Chenevix-Trench G.;
RT "Most germ-line mutations in the nevoid basal cell carcinoma syndrome
RT lead to a premature termination of the PATCHED protein, and no
RT genotype-phenotype correlations are evident.";
RL Am. J. Hum. Genet. 60:21-26(1997).
RN [6]
RN VARIANT NBCCS ASP-1438.
RX MEDLINE=98001068; PubMed=9341860; DOI=10.1007/s004390050541;
RA Lench N.J., Telford E.A.R., High A.S., Markham A.F., Wicking C.,
RA Wainwright B.J.;
RT "Characterisation of human patched germ line mutations in naevoid
RT basal cell carcinoma syndrome.";
RL Hum. Genet. 100:497-502(1997).
RN [7]
RN VARIANT PRO-1315.
RX MEDLINE=99211211; PubMed=10200051;
RX DOI=10.1002/(SICI)1098-1004(1998)11:6<480::AID-HUMU9>3.3.CO;2-W;
RA Hasenpusch-Thell K., Bataille V., Laehdetie J., Obermayr F.,
RA Sampson J.R., Fritschau A.-M.;
RT "Gorlin syndrome: identification of 4 novel germ-line mutations of the
RT human patched (PTCH) gene.";
RL Hum. Mutat. 11:480-480(1998).
RN [8]
RN VARIANTS BCNS SER-376 AND VAL-1083 INS, AND VARIANT BCC TRP-1114.
RX MEDLINE=98281604; PubMed=9620294;
RX DOI=10.1046/j.1523-1747.1998.00222.x;
RA Aszterbaum M., Rothman A.L., Johnson R.L., Fisher M., Xie J.,
RA Bonifas J.M., Zhang X., Scott M.P., Epstein E.H. Jr.;
RT "Identification of mutations in the human PATCHED gene in sporadic
RT basal cell carcinomas and in patients with the basal cell nevus
RT syndrome.";
RL J. Invest. Dermatol. 110:885-888(1998).
RN [9]
RN VARIANT PRO-1315.
RX MEDLINE=20334946; PubMed=10874314;
RX DOI=10.1002/1098-1004(200007)16:1<89::AID-HUMU18>3.3.CO;2-Z;
RA Dong J., Gailani M.R., Pomeroy S.L., Reardon D., Bale A.E.;
RT "Identification of PATCHED mutations in medulloblastomas by direct
RT sequencing.";
RL Hum. Mutat. 16:89-90(2000).
RN [10]
RN VARIANT BCNS PRO-1132.
RX MEDLINE=21154106; PubMed=11231326;
RX DOI=10.1046/j.1523-1747.2001.01279-2.x;
RA Reifemberger J., Arnold N., Klechle M., Reifemberger G., Hauschild A.;
RT "Coincident PTCH and BRCA1 germline mutations in a patient with nevoid
RT basal cell carcinoma syndrome and familial breast cancer.";
RL J. Invest. Dermatol. 116:472-474(2001).
RN [11]
RN VARIANTS SQUAMOUS CELL CARCINOMA MET-829 AND LYS-1242.
RX MEDLINE=21184537; PubMed=11286632;
RX DOI=10.1046/j.1523-1747.2001.01301.x;
RA Ping X.L., Ratner D., Zhang H., Wu X.L., Zhang M.J., Chen F.F.,
RA Silvers D.N., Peacocke M., Tsou H.C.;
RT "PTCH mutations in squamous cell carcinoma of the skin.";
RL J. Invest. Dermatol. 116:614-616(2001).
RN [12]
RN VARIANTS BCNS PRO-230 AND 505-LEU-ARG-506.
RX PubMed=15459969; DOI=10.1002/humu.9289;
RA Savino M., d'Apolito M., Formica V., Baorda F., Mari F., Renieri A.,
RA Carabba E., Tarantino E., Andreucci E., Belli S., Lo Muzio L.,
RA Dallapiccola B., Zelante L., Savoia A.;
RT "Spectrum of PTCH mutations in Italian nevoid basal cell-carcinoma
RT syndrome patients: identification of thirteen novel alleles.";
RL Hum. Mutat. 24:441-441(2004).
CC -!- FUNCTION: Acts as a receptor for sonic hedgehog (SHH), indian
CC hedgehog (IHH) and desert hedgehog (DHH). Associates with the
CC smoothened protein (SMO) to transduce the hedgehog's proteins
CC signal. Seems to have a tumor suppressor function, as inactivation
CC of this protein is probably a necessary, if not sufficient step
CC for tumorigenesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: In the adult, expressed in brain, lung, liver,
CC heart, placenta, skeletal muscle, pancreas and kidney. Expressed
CC in tumor cells but not in normal skin.
CC -!- DEVELOPMENTAL STAGE: In the embryo, found in all major target
CC tissues of sonic hedgehog, such as the ventral neural tube,
CC somites, and tissues surrounding the zone of polarizing activity
CC of the limb bud.
CC -!- PTM: Glycosylation is necessary for SHH binding (By similarity).
CC -!- DISEASE: Defects in PTCH are probably the cause of basal cell
CC nevus syndrome (BCNS) [MIM:109400]; also known as Gorlin syndrome
CC or Gorlin-Goltz syndrome. BCNS is an autosomal dominant disease
CC characterized by nevoid basal cell carcinomas (NBCCS) and
CC developmental abnormalities such as rib and craniofacial
CC alterations, polydactyly, syndactyly, and spina bifida. In
CC addition, the patients suffer from a multitude of tumors like
CC basal cell carcinomas (BCC), fibromas of the ovaries and heart,
CC cysts of the skin, jaws and mesentery, as well as medulloblastomas
CC and meningiomas. PTCH is also mutated in squamous cell carcinoma
CC (SCC). Could also be associated with large body size observed in
CC BCNS patients.
CC -!- DISEASE: Defects in PTCH are a cause of sporadic basal cell
CC carcinoma (BCC) [MIM:605462].
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC -!- DATABASE: NAME=PTCH mutation database;
CC WWW="http://www.cybergene.se/PTCH/ptchbase.html".
CC -!- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocancer/Genes/PTCH100.html".
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL; U59464; AAC50550.1; -; mRNA.
CC EMBL; AL161729; CAH73817.1; -; Genomic_DNA.
DR

DR	GO; GO:0007224; P:smoothed signaling pathway; IGI.	
DR	InterPro; IPR003392; Patched.	
DR	InterPro; IPR004766; Patchedtm_recept.	
DR	InterPro; IPR000731; SSD 5TM.	
DR	Pfam; PF02460; Patched; 1.	
DR	TIGRFAMs; TIGR00918; 2A060602; 1.	
DR	PROSITE; PS50156; SSD; 1.	
KW	Glycoprotein; Receptor; Transmembrane.	
FT	TOPO_DOM 1 86	
FT	TRANSMEM 87 107	
FT	TOPO_DOM 108 422	
FT	TRANSMEM 423 443	
FT	TOPO_DOM 444 458	
FT	TRANSMEM 459 479	
FT	TOPO_DOM 480 487	
FT	TRANSMEM 488 508	
FT	TOPO_DOM 509 533	
FT	TRANSMEM 534 554	
FT	TOPO_DOM 555 563	
FT	TRANSMEM 564 584	
FT	TOPO_DOM 585 734	
FT	TRANSMEM 735 755	
FT	TOPO_DOM 756 1013	
FT	TRANSMEM 1014 1034	
FT	TOPO_DOM 1035 1039	
FT	TRANSMEM 1040 1060	
FT	TOPO_DOM 1061 1069	
FT	TRANSMEM 1070 1090	
FT	TOPO_DOM 1091 1107	
FT	TRANSMEM 1108 1128	
FT	TOPO_DOM 1129 1140	
FT	TRANSMEM 1141 1161	
FT	TOPO_DOM 1162 1434	
FT	DOMAIN 424 584	
FT	CARBOHYD 127 127	
FT	CARBOHYD 298 298	
FT	CARBOHYD 335 335	
FT	CARBOHYD 400 400	
FT	CARBOHYD 861 861	
FT	CARBOHYD 986 986	
SQ	SEQUENCE 1434 AA; 159273 MW; A5E9189E633173D0 CRC64;	
Query Match		
Best Local Similarity 22.9%; Pred. No. 2.2e-36;		
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
QY	356 LVLSVIPVVALAAGLVFTELTDPELVSAAPSQARSEKAFHQHFGPFRTN-QVILTA 414	
DB	88 LVVGLLI FGAFAVGLKAANLETVBELWVEVGRVSRLENYTRQKIGEEAMFNPQLMIQT 147	
QY	415 PNRSSRYDSLLGPKNFGSILDLLELLE--LQERLRHLQVMSPEAQRNISLQDICY 472	
DB	148 PKEEG-----ANVLTEALLQHLDSALQASRVHYVMYN-----ROWKLEHLCY 190	
QY	473 AP-----LNPNDTSYDCCINSLLQVFNQNRLLTLLLTANQTLMGQ----- 512	
DB	191 KSGELITETGYMDQIIEYLYPCLIITPLDCFEGAKLQSGTA--YLLGKPLPLRWTFDPL 248	
QY	513 -----TSQVDWKDHFPLY-----CAN-----APL----- 530	
DB	249 EFLEELKKINYQVDSWEEMLNKAEGVGYMDRPNLNPADPCDPATAPNKNSTKPLDVALV 308	
QY	531 -----TFKQGT-----ALALSCMADYGAVPFPFLAIGYKKG 562	
DB	309 LNGCGOGLSRKYMHOSELIVGTVKNAOKLVSAHALQTMFQMTPKQWYEHFRGY--- 365	
QY	563 DYSEAEALIMTFSLNNYPAGDPPLAQAKLWEEAFLEEMRAFORRMAGMFQVTTAERSLE 622	
DB	366 DY-----VSHINWNE---DRAAAILEAMQRTYVEVHQSVAPNSQKVLPTT-TTLTD 414	
QY	623 DEINRTTAEDLPATSYIIVFLYISLALGYSWSRWMDVSKATLGLGVAIVLGVNVA 682	
DB	415 DILKSFSDSVIRVASGYLLMLAYACLTM---LRWD---CSKSQAGVLAGVLLVALSVAA 469	
QY	683 AMGFPSVLGIRSSILVILQVVPFLVLSVGADNIIFFVLEVQRLPRRPPCEPREVHIGRALGR 742	
DB	470 GLGLCLIGISFNAATQVLPFLALGVDDVFLLAHAFSETQNKRIPEPDRDTGECLKR 529	
QY	743 VAPSMILCSLSEACIFPLGALTMPAVRTPALTSGLAVLDFLLQMSAFVALLSLDSKQQ 802	
DB	530 TGASVALTSISNVTAFMAALIPALRAFSLQAQVVVVVFAMVLLIFAILLSMDLYRR 589	
QY	803 EASRLDVCCC-----VKPOEL-----PPGQGEGL----- 827	
DB	590 EDRLDIFCCFTSPCVSRVIOVEPOAYTEPHSNTRYSPPPPTSHSFAHETHITMQSTVQ 649	
QY	828 ----- 827	
DB	650 LRTEYDPHTHYVYTTAEPSEISVQPVTVTQDNLSQSPSTSTSDLLSQFSDSSLHCL 709	
QY	828 -----LLGPFQKAYAPFLHWTGCVLLFLALFGVLSLYSMCHISVGLDELALPK 879	
DB	710 EPPCTKWTLSSFAEKHYAPFLKPKAKVWVILLFGLLVSLVGTTRVRDGLDLTDIVPR 769	
QY	880 DSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSEAGNAICSSAGCNFSFSTQKIQAT- 938	
DB	770 ETRYDFIAAQFYFSF-----YNM-----YIVTQADVPNI 801	
QY	939 -----EPPEOSYLAIPASS-----WVDDDFIDWL-----TPSS--- 965	
DB	802 QHLLYDLHLKSFNVKVMLEENKQLPQWMLHYFEDMLQGLQDAPDSDWETGRIMPNNYKN 861	
QY	966 -----CCRLYISGPNKDFCPSTVNSLNCIKMCSITWGSVPSVPSVEQFHKL-PWFLN 1017	
DB	862 GSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVSN 915	
QY	1018 D-----RPNIKCPKGLAAY--STSVNLTSDGOVLASRFMAYHKPLKNSQD 1061	
DB	916 DFVAYASQANIRPHREWEHDK---ADYMPETRLRIPAAEPYEAQPPFYVINGLRDTS 972	
QY	1062 YTEALRAARELAANITA-DLRFKVPETDPAPFVFFYTTNVFYEQVLTILPGLFMLSJCL 1120	
DB	973 FVEALEKRVKVCNNVTSGLSSYPNG-----YFP-----LFWEQVLSLRHLLLSISVVL 1022	
QY	1121 VPTFAVSCLLGLDLRSGLLNLLSVMLLVDTVGFMAWDISYNAVLSLINLSVAGMSVE 1180	
DB	1023 ACTFLVCAVFLNFWPTAGII-VVVALMTVELFGMGLIGIKLSAVPVVILLASVIGIVE 1081	
QY	1181 FVSHITRSF--AISTKPTWLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLIQIF 1238	
DB	1082 FTVHVALAFLTAIGDKN---HRAMLALEHMFAPVLDG-AVSTLLGLVLMAGSERPDFIVRY 1137	
QY	1239 FFRMLMLTLGLLHGLVFLPVLSYVG--PDVNPALALEQ---KRAEEVAANVAVSCP 1293	
DB	1138 FFAVLAILTVLGLVGLVLLPVLSFFGCPCEVSPANGNLRLPTSPSPPPSVVRFAVPP 1197	
QY	1294 NHPSRVS-TADNIYNHSEFGSI-----KGAG 1319	
DB	1198 GHTNNGSDSDSEYSSQTTVSGISEELROYEAQQGAG 1234	
RESULT 64		
QSRJ13 BRARE		
ID	QSRJ13_BRARE PRELIMINARY; PRT; 827 AA.	
AC	QSRJ13;	
DT	01-FEB-2005 (Tremblrel. 29, Created)	
DT	01-FEB-2005 (Tremblrel. 29, Last sequence update)	
DT	01-FEB-2005 (Tremblrel. 29, Last annotation update)	
DE	Novel protein (Fragment).	
GN	Name=OTDARP0000008780; ORFNames=DKEY-61p9.10-001;	
OS	Brachydanio rerio (Zebrafish) (Danio rerio).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariphiysi; Cypriniformes;	
OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7955;	
RN	[1]	


```

RP NUCLEOTIDE SEQUENCE.
RA Pelan S.;
RL Submitted (DEC-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BX005238; CAI21193.1; -; Genomic DNA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0008158; F: hedgehog receptor activity; IEA.
DR InterPro; IPR003439; ABC_transp_like.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER_1; UNKNOWN_1.
DR PROSITE; PS0156; SSD; 1.
FT NON_TER 827 827
SQ SEQUENCE 827 AA; 92019 MW; A7F10494A2C9C041 CRC64;

Query Match          9.2%; Score 636.5; DB 2; Length 827;
Best Local Similarity 24.1%; Pred. No. 5.4e-36;
Matches 233; Conservative 175; Mismatches 371; Indels 187; Gaps 33;

QY 337 LQPFQGGCTWASWPLTILVLSVIPVALAAGLVF-TELTTDPVELMSAP-NSQARSEK 394
Db 12 LSLVPEKLGRLGRHPVVFIILSLVYAAGLGAGFIFLKEREANDIEDQFTPVNGPAKDR 71
QY 395 AFHQHGFGRFTNQVILTAHRSYRYSLLGPNFSGIILDLLELLELQERLRL 454
Db 72 EIVAEHFPFSDRFSQRL-----MSEGTVASLIITDQGENILTAFAFEILALDRQVKTL 127
QY 455 QVMSPEAQRNLSQDICVAPLNDPNTSLVDCCINSLLQYFQNNR-----TILLLTANQT 508
Db 128 Q-----HLGNTFEKLC-AKIG-----NCVSNVAVLDIIRYNAADISVTIYPINDKT 174
QY 509 LMGQTSQVMDKDFLYCANAPLTPKDGTAALSLCMADYGAPVFPFLAIGYRGKQYS--- 565
Db 175 FLGTT-----IGVETQPNSSML 192
QY 566 -BAEALIMTFSLNTPAGDPRLAQAQL-WEBAFLEBMAF--ORRMAGHGFQVTFIAERSL 621
Db 193 KSAKAIRLYYFL-----DEKSKGNADWLEGIQIFFSNYTDQEKVS-----VSYFTSVSR 242
QY 622 EDENRTAEDLPFATSVIVIFLISLALGSYSSMSRWMDSKATLIGGVAVVLGAVM 681
Db 243 QNEFEGNTDSVPLFSITTA-----LAINIAV--LSCLRLDCVTKVWALLGVVVSAGMAVL 297
QY 682 AAMGFPSVILGRSSVILQVVPFLVSLVGADNIFIVLEYQRLPRRGPPEVHIGRALG 741
Db 298 ASFGILLFCGMPFAMTV-GSAFLLILGVGVDDMFIMISSWQKTAVDKGV--EFRLAEAYK 354
QY 742 RVAPSMLLCSLSEAICFFIGALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSLDSKR 801
Db 355 EAGVSITITTLTDVLAFAFYIGLLTPFRSVQSFQWYTSALLFCYLFNITPFGACALNGRR 414
QY 802 QEASRLDVCCVKKQELPPGQGGELLG-----FFQKAVAPELLHWI 844
Db 415 EKSNRHFLTCMTVPK---PSGDVAVSCCAGGAFDENTKEHEMPMEVFYKGYGFLAKVM 471
QY 845 TRGVLLLFALFGVSLYSMSCHISVGLDOELALPKDSYLLDYFLFLNRFYEVGAPVYFT 904
Db 472 VKVIVCLIVAGVLAISYICCFQWEGEGLDLKHLATGVSVDYDREDFSFAGFNWMLV 531
QY 905 TLGVNPF---SSBAGNNA-ICSSAGACNPFSTQIKIYATEFPFQSYLAIPASSVDDPFDIM 960
Db 532 INDEHFQVWSPARKSLDLC-----LKNFRDLTMVD-----SEIPLTSLWLDAYM-- 575
QY 961 LTFPSCCRLYISGPNKDFCPSVNSLNCNKMSTMGVSVPSEVQPHKYLFWFLNDRP 1020
Db 576 -----KFGQSSFD-----LNNEMI-----PKTQLPAFIN--- 599
QY 1021 NIKCPKGLAAYSTSVNLTSQCVLASRFMAYHKPLKNSQDYTEALRAARELAENITADL 1080
Db 600 -----RSEFSDVHFT--DNNINATRMFIQTVIKTAIDKMDLNAFRE-AANTCGRLL 649
QY 1081 RKVPGTDPAFEVFPYITNVFYEQYLTILPEGLFMLSLCLVFPFAVSCLLGLDLRSGLL 1140

Db 650 -----ETPVDLIVYHPAPIYFDQYAVIVSNNTQNLVAATCVMLVISLLIHPPLCS-LW 702
QY 1141 NLLSTVMIILVDTVGPMALWDISYNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLSR 1200
Db 703 VTFSTASIVVGLAGFPMALWDISLDSVSNLNVICIGFSDVFSAHISYAFVSSEKSSANEK 762
QY 1201 AKBATISMGSAVFAGVAMTNLPGILVLGLAKAQLQIOPFFRNLNLLITLGLLHGLVFLPV 1260
Db 763 ATDALHKLGYPIIQG-AVSTIAGVVVLAARAKSYIPRT-FFKIMFLVILFGALHGVFLPV 820
QY 1261 ILSYVG 1266
Db 821 FLSPFG 826

RESULT 65
PTCL CHICK
ID PTCL_CHICK STANDARD; PRT; 1442 AA.
AC Q90693;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Patched protein homolog 1 (PTCL) (PTC).
GN Name=PTCH; Synonyms=PTC;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus
OC NCBI_TaxID=9031;
OX [1]
RN [1]
NUCLEOTIDE SEQUENCE.
RP TISSUE=Limb bud;
RC MEDLINE=96205046; PubMed=8620849;
RA Marigo V., Scott M.P., Johnson R.L., Goodrich L.V., Tabin C.J.;
RT "Conservation in hedgehog signaling: induction of a chicken patched
homolog by Sonic hedgehog in the developing limb.";
RL Nature 384:176-179 (1996).
RN [2]
RP CHARACTERIZATION.
RC MEDLINE=97064175; PubMed=8906794; DOI=10.1038/384176a0;
RA Marigo V., Davey R.A., Zuo Y., Cunningham J.M., Tabin C.J.;
RT "Biochemical evidence that patched is the Hedgehog receptor.";
RL Nature 384:176-179 (1996).
CC -!- FUNCTION: Acts as a receptor for sonic hedgehog (SHH), indian
hedgehog (IHH) and desert hedgehog (DHH). Associates with the
smoothed protein (SMO) to transduce the hedgehog's proteins
signal.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expression is seen in the embryonic neural
tube, sclerotome, visceral mesoderm, and limb bud.
CC -!- DEVELOPMENTAL STAGE: In stage 10 embryo, expression is seen in
neural tube, and at lower levels in the notochord, epithelial
somites, endoderm and splanchnic mesoderm. At stage 18, PTC is
broadly expressed in the neural tube but excluded from the cells
of the floor plate. At stage 32, expression occurs in the
mesodermal cells of the gastrointestinal tract.
CC -!- INDUCTION: Activated by hedgehog; repressed by itself (Probable).
CC -!- PTM: Glycosylation is necessary for SHH binding.
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use as long as its content is in no way modified and this statement is not
removed.
CC EMBL; U40074; AAC59898.1; -; mRNA.
CC PIR; T18538; T18538.
CC Ensembl; ENSGALG00000012620; Gallus gallus.
CC InterPro; IPR003392; Patched.
CC InterPro; IPR004766; Patchedtm_recept.

```

DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; 1.
 DR TIGRFAMs; TIGR00918; 2A060602; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Glycoprotein; Receptor; Transmembrane.
 FT TOPO DOM 1 101
 FT TRANSMEM 102 122
 FT TOPO DOM 123 436
 FT TRANSMEM 437 457
 FT TOPO DOM 458 472
 FT TRANSMEM 473 493
 FT TOPO DOM 494 501
 FT TRANSMEM 502 522
 FT TOPO DOM 523 547
 FT TRANSMEM 548 568
 FT TOPO DOM 569 577
 FT TRANSMEM 578 598
 FT TOPO DOM 599 747
 FT TRANSMEM 748 768
 FT TOPO DOM 769 1026
 FT TRANSMEM 1027 1047
 FT TOPO DOM 1048 1053
 FT TRANSMEM 1054 1074
 FT TOPO DOM 1075 1082
 FT TRANSMEM 1083 1101
 FT TOPO DOM 1102 1120
 FT TRANSMEM 1121 1141
 FT TOPO DOM 1142 1153
 FT TRANSMEM 1154 1174
 FT TOPO DOM 1175 1442
 FT DOMAIN 438 598
 FT CARBOHYD 141 141
 FT CARBOHYD 312 312
 FT CARBOHYD 349 349
 FT CARBOHYD 414 414
 FT CARBOHYD 827 827
 FT CARBOHYD 874 874
 FT CARBOHYD 999 999
 SQ SEQUENCE 1442 AA; 160578 MW; 97355F17FB886843 CRC64;

Query Match 9.1%; Score 628.5; DB 1; Length 1442;
 Best Local Similarity 23.6%; Pred. No. 4.2e-35;
 Matches 284; Conservative 165; Mismatches 421; Indels 335; Gaps 47;

QY 365 ALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHGFPPFTN-QVILTAENRSRYD 423
 111 AFVAGLRAANLETNVEELVVEGVRSRELNTYRQKIGSEAMFNQMLTIQTPQEDG--- 166
 QY 424 SLLLGPKNFGSLDLLELLE--LQERLRHLQVMSPEAQNISLQDICYAP----- 474
 167 -----TNVLTTEALRQHLDSALQASRVHYMYN---RQWKLEHLCKYKSGELITEA 213
 QY 475 --LNPDTSLYDCCLNSLIQYFQNNRTLLLTANOTLMGQ----- 512
 214 GYMDQIIIEVLYPCLIIITPLDCFWEGAKLQSGTA--YLLGKPLQWINFDPDLFLBELKKI 271
 QY 513 TSQVDKWDHLY-----CAN-----APLTF--KQGT-----ALALS--CMADY 546
 272 NYQVESWEEMLNKAERVGHGYMDRPNCLNPADPCFIPAPKNKSTKPLDVALVLSGGC---Y 328
 QY 547 GAPVFPF-----LAIGG-----YKGXDYSAEALIMTFSL-----NNYP----- 580
 329 GLSRKYWHQBELIIIGTVKNSGKLVS-AQALQTMFQMLTPKQWYHEFKGYEYVSHINW 387
 QY 581 AGDPRLAQAKLWEAFLEENKRAFORRMAGNFQVTTFAERSLEDEINRTTAEDLPATSY 640
 388 NEDKAAAILAEWQRMVYEVVHQSVQNSQKLSFTT-TLDDILKSFSDSVIRVASGY 446
 QY 641 IVFLYISLALGYSYSSNSRWVDSKATLGLGGVAVVLGAVMAAGPFYSILGRSSVLILQ 700
 447 LLMLAYACLTM---LRWD--CAKSGAGVAGLAVLLVALSVAAAGLGLCSLIGISFNATQ 501
 QY 701 VVPLVLSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEACFFL 760

Db 502 VLPFLALGVGVDDVFLLAHAFSETQNKRIPEFDRTEGCEKRTGASVALTSISNVTAFPM 561
 QY 761 GALTMPAVRTFALTSLGLVILDFLOMSAFALLSLDSKQASRLDVCCC----- 812
 Db 562 AALIPALRAFSLQAAVVVVFNFAMVLLIFPAILSMOLYRRDRRLDIFCCFTSPCVTR 621
 QY 813 ---VKPQ-----ELPPPGQGGGL----- 827
 Db 622 VIQIEQAYAEENDNICYSPPPYSSHSFAHETQITWQSTVOLRTEYDPHTQAYVTTAEPR 681
 QY 828 -----LLGFFQKAYAP 838
 Db 682 SEISVQPVTVTQDSLSCQSPESASSTRDLLSQFSDSVHCLPEPPCTKWTLSFAEKHYAP 741
 QY 839 FLLHWITRGVLLFLALPGVSLYSMCHISVGLDQELALPKDSVLLDYELFLNRYPEVGA 898
 Db 742 FLLKPKAKVVVIFLGLGLSLYGTTRVRDGLDLDIVPRDREYDFIAAQPKYFSF-- 799
 QY 899 PVYFVTTLGVNFSSEAGMNAICSSAGCNFSTFKIQIYAT-----EFPQSYLAI 948
 Db 800 -----YNM-----YIVTKADYPNVQHLLEYELHRSFSNVTVLL 833
 QY 949 PA-----SSWDDPIDWL-----TPSS-----CCELYISGNKD 977
 Db 834 EGDRLPKMWLHYPRDWLQLODAFSDSWETGKITYSNYKNGSDDAVLAYKLLVQTGNRA 893
 QY 978 KFCSTVNSLCLNKNCSITWGSVRPSVEOPHKYL-PWFLND-----RPN1 1022
 Db 894 K--PIDISQLT--KQRLVDADGIINPA--FYIYLTAVSNDPVAYAASQANIRPHPEW 947
 QY 1023 KCPKGGGLAA--STSVNLTSQGVLASRFMAYHKPKNSQDYTBALRAARELANITA-D 1079
 Db 948 VHDK--ADYMPETRLRAIPAEPIEYAPFPYNGLRSETSDFEAEIKVRAICNNYISLG 1004
 QY 1080 LRKVPDTPAPEVPPYITNVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGL 1139
 Db 1005 IASYPNG-----YFP-----LFWEQYIGLRHMLLSISVVLACTFLVCALFLNPNWTAGI 1054
 QY 1140 LNLISIVMLVDTVGFMAWMDISYNAVSLINLAVGMSVEFVSHITRSP--AISTKPTW 1197
 Db 1055 I-VVVALMTVELFGMGLIGIKLSNPVVLIIASVIGVEFTVHIALAFLTAIGDKN-- 1111
 QY 1198 LERAKETISMGSAVFAGVAMTNLPGILVLGLAKAQIPIFFRLNLILITLLGLHLGVF 1257
 Db 1112 -RRVLALEHMFAPVLDG-AVSTLLGLVLMAGSEFDFIVRYFFAVLAITLITGLVNLGLV 1169
 QY 1258 LPVILSYVG--PDVNPALA---LEQKRAEEAAMVAVASCPNHPRSVSTADNIYVNHSPF 1312
 Db 1170 LPVLLSFFGYPPEVPSPACGRNRLPTPSPPPSIVRFALPPGHTNNGS--DSSDSEYSQ 1227
 QY 1313 GSIKG 1317
 Db 1228 TVVSG 1232

RESULT 66
 QSVZC2 HUMAN
 ID QSVZC2_HUMAN PRELIMINARY; PRT; 1296 AA.
 AC QSVZC2;
 DT 01-FEB-2005 (TrEMBLrel. 29, Created)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
 DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
 DE Patched homolog (Drosophila).
 GN Name=PTCH; ORFNames=RP11-43505.3-002;
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Clark G.;

Db 321 SKYVHMQBELII---GGTVKNASQOI-----VSALALQTMFQLWTP 359
Qy 550 VFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPLRLAQAKLWBEAFLEEMRPFQRRMAG 609
Db 360 KQMY-----EHFKGHEVVSHWNNE-----DKAAALAEAWQRTYVQVHQSVQPNSS 406
Qy 610 MFQVTFATERSLEDEINTTADLPFATSYIVIFLYTSLALGSSYSSKSRVWVDSKATLG 669
Db 407 QKVLFPFTT--TTLDLILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CAKSGQAVG 460
Qy 670 LGGVAVLGVAAWMAWPFYSYLGIRSSVILQWPFVLVSVGADNIFIFVLEYQRLPRRP 729
Db 461 LAGVLLVALSVAAGLGLSLGISFNATTVQLPFLALGVGVDDVFLLAHAFSETGQNK 520
Qy 730 EPREVHIGRALGRVAPSMLLCSLSAICFFLGALTTPMPAVRTFALTSLAVILDFLOMS 789
Db 521 IPFEDRTGECLKRTGASVALTSISNTAFFMAALIPALRAFSLQAAVVVFNFAVLL 580
Qy 790 AFVALLSDSKRQEARLDVCCC-----VKPQ-----ELPP----- 820
Db 581 IPFALLSMDLYRREVRLDIFCCFSPCVSRVQIEPOAYTDNDNTNNTKYSLPPTVSSHSP 640
Qy 821 -----PGQ----- 823
Db 641 AHETOITMQSTVOLRTEYDPRQLYYTTAQRSEISVQPAASTPDQVSGQTPESTSTRD 700
Qy 824 -----GEGLLGFFQKAYAPFLLHWITRGVVLVLLFALFGVSLYSWCH 866
Db 701 LISQFSVHGGSMQCTPDSKWTLSSEAFAKHYAPFLPKTKKAVILGLFALLSVSLYGTTR 760
Qy 867 ISVGLDQELAPKDSYLDYFLFNRYFEGVAPVYFVITLGNFSSAGMAICSSACGN 926
Db 761 VRDGLDLTDIVPRETREYDFIATQKYSF-YHMYVVVTKA----- 800
Qy 927 NFSFTQKIOYA--TEFPEOSYLAI PASS-----WDDFDIMLTPSS----- 965
Db 801 DYPRAQLLYELHKEFVGVYVLLSGNKQLPKWMLHYFRDMLQGLQDTFDHEWEAGKITR 860
Qy 966 -----CRLYISGPNKDFCPSTVNSLNCNLSITMGSRVPSVEQHKYL- 1012
Db 861 NDRNASDDAVLAYKLLIQTGNSDK--PINLQLT--KQRLVDADGIIQPN--FYIYLT 914
Qy 1013 PWFILDRPNKICPKGLAAYTSV-----VAIASQANIRPHPEWHLKADRPETRTIRAAPIEYVPP 1050
Db 915 AWSNDP-----VAIASQANIRPHPEWHLKADRPETRTIRAAPIEYVPP 964
Qy 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYTTITNVFQEYLTIIP 1110
Db 965 FYINGLRBTSDFEAELEKRAICNNVTS-----LGVSYPNGYPP-----LFWEQVLSRH 1015
Qy 1111 EGLFMLSCLVPTTFVAVSCILLGLDRLSGLLNLSIVMILVDTVGFMAWLDISYNAVSLIN 1170
Db 1016 WLLSISVVLACTFLVLCALFLNPNPTAGII-VVVALMTVELFGMMGLIGIKLSAPVVI 1074
Qy 1171 LYSVAGMSVEFVSHITRSP--AISTKPTWLERAKEATISMSASVAFVAVMTNLPGLVLG 1228
Db 1075 LTASVGIGVEFVHVALAFLTAVGDKN---RRAVLAHEMFAPVLDG-AVSTLLGVLMIA 1130
Qy 1229 LAKAQIQLIFFFRNLNLTLLGLLHGLFPLVILSYVG--PDVNPA-----LALQKRAE 1281
Db 1131 GSEFDFIVRYFFFAVAILLILGVNLGLVLLPVLSPFGYPYEPVSPITNGDRLVSDPPP 1190
Qy 1282 EAVAAMVASCNPHRSVST-ADNIYVNHSPFSGIKG 1317
Db 1191 N-----IVRPEHPRPTQNGSDSDSEYSSQTSVSG 1221

RESULT 68
Q98SW6_XENLA
ID Q98SW6_XENLA PRELIMINARY; PRT: 1418 AA.
AC Q98SW6;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Patched-1.
GN Name=Ptcl;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=21095181; PubMed=11165486; DOI=10.1016/S0925-4773(00)00526-8;
RA Koebnick K., Hollemann T., Pieler T.;
RT "Molecular cloning and expression analysis of the Hedgehog receptors
XPtcl and Xsmo in Xenopus laevis.";
RL Mech. Dev. 100:303-308(2001).
DR EMBL; AF302765; AAL5463.1; --; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR00731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
SQ SEQUENCE 1418 AA; 158226 MW; 4FEA30C2D5031F54 CRC64;
Query Match 8.9%; Score 617.5; DB 2; Length 1418;
Best Local Similarity 21.6%; Pred. No. 2.5e-34;
Matches 280; Conservative 184; Mismatches 452; Indels 381; Gaps 43;
Qy 306 RVAPARDKSNVDPKGTSLSKLSFSFTHTLGQFFQG-----W----- 344
Db 21 RVRRRRGRSRVAPPDHDYLRQPSYCDANFALQOISEGAKRGKAPLWLRFAFRQLFKL 80
Qy 345 GTWVASPLTILVLSVIPVVALAAGLVFTELTDPVELWSAPNSQARSEKAPHQHFQ-- 402
Db 81 GCYIQKCGFLVVGLLIFGAFVGLRAANLETNVEELWVEVGGVRSRELDYTRQKIGEE 140
Qy 403 -----PF-----PFTNQVI 411
Db 141 AMFNPLMIOTPLEDGANVLTTEALLQHLHSALEATKVQVVMYKNPKWLELCFKSGELI 200
Qy 412 LTAPNRSY---RYDSLIL-----GPKNFSGILD---DLL-----LELL-ELQE 449
Db 201 TEAVVSVQIISMTPLIITPLDCEWEGAKLQSGWYLPQKDIQWTFDPLELLEELKX 260
Qy 450 RLRLHQVWSPQAQR-----NISQDICYAPLNPDNTSLYDC-----CINS 489
Db 261 GKLDHDIWEEMINKAEVGHGYMDRPLNPSDKNCPYAPNKNSTKPDVSVLSLGGCYGL 320
Qy 490 LLOVFNQNRULLLTANOTLMGQTSQVDWKHFLYCANAPLTFKDGTAALALSCMADYGP 549
Db 321 SKKYMHMQBELII---GGTVKNASQOI-----VSALALQTMFQLWTP 359
Qy 550 VFPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPLRLAQAKLWBEAFLEEMRPFQRRMAG 609
Db 360 KQMY-----EHFKGHEVVSHWNNE-----DKAAALAEAWQRTYVQVHQSVQPNSS 406
Qy 610 MFQVTFATERSLEDEINTTADLPFATSYIVIFLYTSLALGSSYSSKSRVWVDSKATLG 669
Db 407 QKVLFPFTT--TTLDLILKSFSDSVIRVASGYLLMLAYACLTM---LRWD--CAKSGQAVG 460
Qy 670 LGGVAVLGVAAWMAWPFYSYLGIRSSVILQWPFVLVSVGADNIFIFVLEYQRLPRRP 729
Db 461 LAGVLLVALSVAAGLGLSLGISFNATTVQLPFLALGVGVDDVFLLAHAFSETGQNK 520
Qy 730 EPREVHIGRALGRVAPSMLLCSLSAICFFLGALTTPMPAVRTFALTSLAVILDFLOMS 789
Db 521 IPFEDRTGECLKRTGASVALTSISNTAFFMAALIPALRAFSLQAAVVVFNFAVLL 580
Qy 790 AFVALLSDSKRQEARLDVCCC-----VKPQ-----ELPP----- 820

[illegible]

RESULT 69

```

RESUL 69
Q9W6T6 BRARE PRELIMINARY; PRT; 1243 AA.
ID Q9W6T6 BRARE PRELIMINARY; PRT; 1243 AA.
AC Q9W6T6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched-2 protein.
DN Name=ptc2; Synonyms=ptcl;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Lewis K.E., Concordet J.P., Ingham P.W.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RS EMBL; AJ007742; CAB39726.1; -; mRNA.
DR Ensembl; ENSDARGO000016404; Danio rerio.
DR ZFIN; ZDB-GENE-980526-196; ptc2.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.

```

```
Db 922 VSADGIINPNA--FTYILSAMVSNDFVAYAAASQANIRPH-----PWEHLHRTDTSIPASRLN 976
Qy 1038 LTSQDVQLASRFMAYHKLNQSDYTEALRAARELAANITADLRKVPGTDPAPFVFPYTI 1097
Db 977 IPAAEPIEYAQFPFVLNGLRTPQVEVAIESVRAICNNYSRQ-----GLPSYPNGY 1027
Qy 1098 TNVFEQYLTLPGLFMLSCLVPTFAVSCLLGLDLRSGLLNLLSIVMTLVDTVGEMA 1157
Db 1028 PFLFEQYVGLRHMLLSISVLAETFLVCAVFLNPNWTAGII-VLVLSTMTVELFGMNG 1086
Qy 1158 LWDISYNAVLNLVSAVMSVEFVSHLTRSFALSTKPTWLERAKEATISMGSAVEAGV- 1216
Db 1087 LIGIKLSAVPVVILLTASVIGVEFTVHVALAFLTAIG-----DRNKRAVLAL-EHMFAPVL 1141
Qy 1217 --AMTNLPGILVLGLAKAQLQIOPFRNLMLITLGLHLGLVFLFVILSYVVG--PDVNP 1272
Db 1142 DCAFTSLGLVLMAGSEDFIVRYFFAVLAIIITVLGVLNGLVLLPVLVLSYFGPCPEVSPA 1201
Qy 1273 ---LALQGRABEAAVAAVWVASCNHPHRSRVSTADN 1304
Db 1202 DGRSLRPTSPBPQPQVVRFTMRPSHTTPEGAGSDS 1236

RESULT 70
Q9DEF3 XENLA PRELIMINARY; PRT; 1413 AA.
AC Q9DEF3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched-2.
GN Name=Xptch-2;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus; Xenopus.
OX NCBI_TaxID=8355;
[1]
NP NUCLEOTIDE SEQUENCE.
RX MEDLINE=20500895; PubMed=1104611; DOI=10.1016/S0925-4773(00)00436-6;
RA Takabatake T., Takahashi T.C., Takabatake Y., Yamada K., Ogawa M.,
RA Takehima K.;
RT "Distinct expression of two types of Xenopus Patched genes during
early embryogenesis and hindlimb development.";
RL Mech. Dev. 98:99-104(2000).
DR EMBL; AB037688; BAB18575.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; ZA060602; 1.
DR PROSITE; PS0156; SSD; 1.
SQ SEQUENCE 1413 AA; 157372 MW; 1215D2D7733E7F17 CRC64;

Query Match 8.7%; Score 599.5; DB 2; Length 1413;
Best Local Similarity 22.0%; Pred. No. 4.6e-33;
Matches 258; Conservative 180; Mismatches 390; Indels 347; Gaps 44;

Qy 355 ILVSLVPVVALAAGLVFTELTDPVELMSAPNSQARSEKAFHQHFG-PFFRTNQVILT 413
Db 83 VLFIGLLVFGALAVGLRVASIEDIERLWVEAGSRVSHELRYTKEXLGEESVYTSQMLIQ 142
Qy 414 APNRSRYDLSLLGPNFGSGLDLDLLELLEQLERHLQWSPDAQNLSLODICY- 472
Db 143 TPREG-----ENILTHEAL--LHLRAALAASKVQVSMYKGSMDLNKICYK 187
Qy 473 --APLNPDN-----TSLYDCCNLSLQVFNQNTL-----LLLTANQTL----- 509
```

```
Db 188 SGVPIIENGMIERIEKLPFCVITITPLDCFEWGSQGGSAIPLFGRADIQTWITLNDPLRLM 247
Qy 510 --MGQTSQVDWKDHFLYCANAPLTFKDGTAI-----ALSCWAD- 545
Db 248 BELGQFTSLDGFKEMLDKAEVQYMERPCLDPTDSDCPRESSNNKTKKPNDIVSTLUGG 307
Qy 546 -YG-APVPEP---LAIGG-YKGG--DYSEAEALIMTFSL-----NNYPAGDPR 585
Db 308 CYGFSKFKMTWQKELILGGMKKGNGELKSAEALQTYLLMSPGQLFEHFKDDYIEIHIN 367
Qy 586 LAQAKL-----WEEAFLEEMRAFQRMMAGMFQVTFATRSLEDEINRTTAEDLPIFATS 639
Db 368 WNEEKAAAILLESQKFEVQ-----AQOSVPQNSQDIHAFSTT 406
Qy 640 -----YIVIFLYISALGSSWSRVWDSKATLGLGGVAVVLGA 679
Db 407 TLNDIMKSFSDVSAIRVAGGYLLMLAYACVTL---LRWD--CSKSQGAAGVAGVLLVALS 461
Qy 680 VMAAMGFFSYLGIRSSLVILQVVPFLVLSGADNIFIIVLEYQRLPRRPGEPREHIGRA 739
Db 462 VASGLGLCSLLGTSFNAAATTOVLPFLALGIGVDDMFLLAHAFTETSR--STPFKDTGEC 519
Qy 740 LGRVAPSMLLCSLSEACIFFLGAITPMPAVRTFALTSGLAVIDFLQMSAFVALLSLDS 799
Db 520 LRRTGTSVTLTSTINNMLAFAFMAALVPIPALRAFSLQAAVIVVFNFAVLLIFPAILSLDL 579
Qy 800 KRQEASRLDVCCC-----VKPQE----- 817
Db 580 QRREDQRLDLCCFYSPCSSRVIQIQPQFSDANDNHTYHPSSYGNPTTSTQITTVQ 639
Qy 818 -----LPPPGQ-----GEGE----- 827
Db 640 APTQCDPSGHHIVITLPTTSQISTSPSVIVPTMDPLGSIPTSPSSSTRDLAQLDESKEK 699
Qy 828 -----LLGFPQKAYAPFLHWTIRGTVLLLFALFGVLSLISMCHISVGLDEL 875
Db 700 RECVPFLKWSLSDFAREKYAPVLLKAETKGIIVVTFMALLGLGLGTYTMVHDGLYLT 759
Qy 876 ALPKD-----SYLLDYFLNRYFVGAPVVFVTTILGYNF--SSEAGMNAICSSAGCN 927
Db 760 IVPRETKEYNFISAQKYFSPYN-----MFIVTKDGFDPKQAESLYDLHEAFGSVK 811
Qy 928 FSFTQKIQYATEFPEQSYLAIPASSWVDDFDLWL-----TPSS----- 965
Db 812 YVREE---GRDLPKM-----WLHYFQDMLRGLQTSFDMDWESGVITQDSYRNGTE 859
Qy 966 ----CCRLYI--SGPNKDKPCPSTVNSLNCNKMCHSITMGSVRPSVEQHKYLP-WFLNDR 1019
Db 860 DGVLAYKLLIQTGNKKEPF-----NFNQLTSRRLVDEKGLIPQ-DAFYIYITVWVSND- 911
Qy 1020 PNIKCPKGLAA-----YST---SVNLTSDGQVLASRFMAYHKLNQSD 1061
Db 912 -----PLGYAASQANFYQPPPEWHIDRYDTTGGENLRIRIPATFPIEFAQPFYNGLRQTS 966
Qy 1062 YTEALRAARELAANITADLRKVPGTDPAPFVFPYTIITNVFEQYLTLPGLFMLSCLIV 1121
Db 967 FVEALESVRSVCEEVVKQ-----GVHSYSPGYPFLFEWQYIGLRHFWLLAISVLA 1017
Qy 1122 PTFVAVSCLLLGLDLRSGLLNLLSIVMLVDTVGFMALWDISVNAVLNLVSAVMSVEF 1181
Db 1018 CTFVLCAILLNPNWTAGII-VFILAMTVELFGINGLIGIKLSAIPVWILASVIGVBF 1076
Qy 1182 VSHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVLGLAKAQLQIIF 1239
Db 1077 TVHVALGFLTAIGDRN---QRSVLAEHMFAPVLQD-AISTLLGVLMLAGSEDFTLRYF 1132
Qy 1240 FRNLNLTLLGLHLGLVFLPILSYVGP--DVNPA 1272
Db 1133 FAVLTILITLGLNGLVLLPVLVLSLIGPPAEVTPA 1167

RESULT 71
Q6IRA5_XENLA
```



```
RESULT 72
P91129 CABEL PRELIMINARY; PRT; 933 AA.
AC P91129;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Patched related family protein 2.
GN Name=ptr-2; ORFNames=C32E8.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RC MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
   investigating biology.";
RL Science 282:2012-2018 (1998).
DR EMBL; U88308; AAB42325.1; -; Genomic_DNA.
DR PIR; T25600; T25600.
DR Ensembl; C32E8.8; Caenorhabditis elegans.
DR WormBase; WBGene0004217; ptr-2.
DR WormPep; C32E8.8; C808532.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR001036; Acrlvin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 933 AA; 104354 MW; C12E5376315FDD11 CRC64;

Query Match      8.7%; Score 598.5; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 3.1e-33;
Matches 251; Conservative 181; Mismatches 362; Indels 303; Gaps 37;

QY 341 FCGGTWASWPLTILVLSVIPVALAAGLVFTELTTPVELWSAPNSQARSEK-AFH-- 397
DB 14 FRLGLICDHPPLFPFVPLPPLTAAMGVGLHLNPLSDAVYLFTPLGAQSKMERMSIHEK 73

QY 398 ----DQHFGP---FFRTNQVILTPNRSYRYSLLGPKNPSGILDLDLLELLELQER 450
DB 74 WPLTDNNYIPGRAVYQSRQEIQTALRN----DSNILDPKFANAVYQLDKY-----IQTR 124

QY 451 LRHLQVSPQAQRNISLQDICVAPLNPNQNTSLYDCINSLLQYFQNNRTLLLTANQTLM 510
DB 125 VRVLH----- 129

QY 511 GQTSQVDKDKHFLYCANAPLTKDGTALA-----LSCMADYG---APVPFLAIGCY- 559
DB 130 -----NGHYYSYKNLCQYKNGCPSNKHVHLSLDLHNGFNITYPYFRFGSEGGYI 181

QY 560 -----KGKDYSE-----AEALIMTFSLNNYP-----AGDPRLAQAKLWEAFLEE 599
DB 182 GSSLGGVTVKGENETDILASAKANFMVHLKFEEMSYISGE-----WE---LEL 230

QY 600 MEAFQRMAGMF-QVTFTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALSGY--SS 656
DB 231 GRMLTQYDPDPYISITFHSQTLADBLKGNADTLIPRFIISITLLIVFSTLCSLFDGS 290

QY 657 WSRVWVDSKATGLGCVAVVIGAVMAAMCFYSLGIRSLVILQVVPFLVLSVGNDFEI 716
DB 291 FSDIWLKSPILSIILGWSAGAITGTGVGLSLMGMPYN-DIVGVMPLVLAVGVNDNML 349

QY 717 FVLEYQRLPRRPGPREVH--IGRALGRVAPSMLLCSLSEACFFLGALTPMPAVRTAL 774
```

```
DB 350 MVAAV----RRTSRTHVHERMGECLEADAANVSLITSDVLSPFGVGAITTPAVQIPCV 405
QY 775 TSGLAVIDFLLQMSAFVALLSLDSKQEAER-----LDVCCCVK----- 814
DB 406 YTGVAIPFAFYQITFFAACILAMKHEASGRNSLFIEAVSABKTSLSLTFQRLFNIGS 465
QY 815 -PQBLPPPGQGGGLLGGFFOKAYAPFLHMTTRGVLLLF-LALFGVSLYSWCHISVGLD 872
DB 466 VPDHSHADVKQPLTSRFFGGEWYAPVLMHPVVRGIAMWVFVYLLGAS-YGCSRIKEGLE 524
QY 873 QELALPKDSYLLDYFLFLNRYF-EVGAPVYV-----TTLGVNFSE 913
DB 525 PVNLLEDVSYAIPHYRLLEKFKYKQGVQVIVINNAPDLRNHTSRDRVHAMVLDPATSKH 584
QY 914 A-GNAICSSAGCNFFSTQKIQIATEPEQ-----SYLAIPASSWVDPDFIWL 961
DB 585 AIGMESV-----QFWLFEMERYYQKLEVEQIIDSFFYGLLHFLASKTNPLAEDIYW- 637
QY 962 TPSSCCRLYISGPNKDKFCPSTVNSLNCNCKMSITMGSVRPSVEQPHKYLFWFLNDRPN 1021
DB 638 -----GPMPPD----- 642
QY 1022 IKCPKGLAAYSTSVNLTSDGOVLAS-RFMAVHKLKNSQDYTEALRAARELANITADL 1080
DB 643 -----DDNGTWNKSPRFILGMDLVTTMDQTDATMSFREVAARW---- 681
QY 1081 RKPFGTDPAPFPYPTTNNVEQYLTLPGLFMLSCLVPTFAVSCLLIGLDRSGLL 1140
DB 682 -----PEFNVTTFMPIWMFTDQYIIIPNTQNIIALLVMIIVIAVLFIQPMPCS-LW 733
QY 1141 NLLSIVMLVDTVGFMALDISYNAVSLINLVSAVMSVEFVSHTRTSFAISTKPTWLER 1200
DB 734 VALACASIDFQVIGVITLNGVNLDAISMITIIMIGSVSDYSAHAYGVVVSREDTAAGR 793
QY 1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQIQLIFFPRLNLLITLLGLLHGLVLPV 1260
DB 794 VKEALSALGWPLSQG-AMSTIIAVSLADIAPYMI-VTFEKTVVLISLGLLHGLVLPV 851
QY 1261 ILSVY-----GPDVNPALALEQKRAEA-----VAAVNVVASCNHPHSRVS 1300
DB 852 LLSIFVRGCCITIPSPGHPSAQKIEQIRIAAIISSSPFLDLRTVAPLRASSPISPHRLE 911
QY 1301 -TADNIYVNHSEFGSIK 1316
DB 912 YTDESPTVNRSKNSIK 928

RESULT 73
Q617L5 CAEBR PRELIMINARY; PRT; 896 AA.
ID Q617L5 CAEBR PRELIMINARY;
AC Q617L5;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG14920.
GN Name=CBG14920;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C. briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
   EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
   preliminary data.
CC EMBL; CAAC01000068; CAB68940.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003392; Patched.
```

DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 896 AA; 100393 MW; 3B6B90BE1FF12EB6 CRC64;

Query Match 8.6%; Score 597; DB 2; Length 896;
Best Local Similarity 23.2%; Pred. No. 3.8e-33;
Matches 232; Conservative 182; Mismatches 378; Indels 206; Gaps 36;

QY 339 QFFQCGTGWASWPLTILVSVIPVVALAAGVFTLTTPVELMSAPNSQARSE-KAFH 397
DB 14 KFFHSYGLLVSGYPFIFLVSPILVTCFLATGLIGLEPLSDALYLYTTPDAPSKTERRIH 73

QY 398 DQHFQGFPTNQVILTAPNRSRYDSSLILGPKNFSGIILDLLELLELQRLRLQVW 457
DB 74 DA--WPLVDGTFVAGRAVTS--REVQVAVVARSGNILLDRVPSNELKLMESFIRN-NIT 128

QY 458 SPEAQRNTSLQDICVAPLNPDNTSLYDCCINSLQYFQNNRTLLLLLTANQTLMGQTSQVD 517
DB 129 VEFNRTWSFPDLCLA--GPDGR-----CANN-----LASRLHQ-- 164

QY 518 WKDHFYLCANAPLTFKDGTALALSCMADYGAPVFPFLAIGGVK-----GKD-YSEAEAL 570
DB 165 ---HGINTYPTVRLSDKSAYIAS-----ALGGVKLAKGNGENIIVEATAW 208

QY 571 INTFSLNNYP-----AGDPRLAQAKLWEEAFLEENRAFORRMAGMFQVFTTFAERSLEDE 624
DB 209 LLIIYQLKFPNEISYVSG-----LWEREFKNQDEY-KKQAKYISITYFHSQTLSD 259

QY 625 INRTTAEDELPIPATSYIVIFY-----ISLALGYSYNSWRVMDSKATLGLGGVAVVIG 678
DB 260 LNRNAERLAPKFIGAFVILVCFSLVCSIVTIKSGYIDW-----VVTKPILSLVGSNAGM 315

QY 679 AVMAWAGPFSYLGIRSSLVILQVPLVSVCAADNIFIPVLEYQRLPRPGPREVHIGR 738
DB 316 GIASAMGMLTYLEIQYN-DIIAVMPFLVAVGTGDNMLVMSLAKRTDR--NLKYDQRIAE 372

QY 1230 AKAQILQIHFPRFLNLLITLLGLLHGLVFLPVILS-YVG 1266
DB 839 VPAYMI-VTFPKTVFLAISIGFLHGLVFLPLMLSVFVG 875

RESULT 74
P91346 CAEBL PRELIMINARY; PRT; 900 AA.
AC P91346;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Patched related family protein 10.
GN Name=ptr-10; ORFNames=F55F8.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Telodermata; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; U80447; AAB37812.3; -, Genomic_DNA.
DR PIR; T29590; T29590.
DR Ensembl; F55F8.1; Caenorhabditis elegans.
DR WormBase; WBGene0004224; F55F8.1.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 900 AA; 100415 MW; 12A301D83326A464 CRC64;

Query Match 8.6%; Score 592; DB 2; Length 900;
Best Local Similarity 23.3%; Pred. No. 8.6e-33;
Matches 237; Conservative 170; Mismatches 369; Indels 242; Gaps 36;

QY 339 QFFQCGTGWASWPLTILVSVIPVVALAAGVFTLTTPVELMSAPNSQARSE-KAFH 397
DB 14 KFFHSYGLLVSGYPFIFLVSPILVTCFLATGLIGLEPLSDALYLYTTPDAPSKTERRIH 73

QY 398 DQHFQGFPTNQVILTAPNRSRYDSSLILGPKNFSGIILDLLELLELQRLRLQVW 457
DB 74 DA--WPLVDGTFVAGRAVTS--REVQVAVVARSGNILLDRVPSNELKLMESFIRN-NIT 128

QY 458 SPEAQRNTSLQDICVAPLNPDNTSLYDCCINSLQYFQNNRTLLLLLTANQTLMGQTSQVD 517
DB 129 VQFSNRTWSFADLCIA--GPDGR-----CANN--HQ-----LASRLHQ-- 164

QY 518 WKDHFYLCANAPLTFKDGTALALSCMADYGAPVFPFLAIGGVK-----GKD-YSEAEAL 570
DB 165 ---HGINTYPTVRLSDKSAYIAS-----ALGGVKLAKGNGENIIVEATAW 208

QY 571 INTFSLNNYP-----AGDPRLAQAKLWEEAFLEENRAFORRMAGMFQVFTTFAERSLEDE 624
DB 209 LLIIYQLKFPNEISYVSG-----LWEREFKNQDEY-KKQAKYISITYFHSQTLSD 259

QY 625 INRTTAEDELPIPATSYIVIFY-----ISLALGYSYNSWRVMDSKATLGLGGVAVVIG 678
DB 260 LNRNAERLAPKFIGAFVILVCFSLVCSIVTIKSGYIDW-----VVTKPILSLVGSNAGM 315

QY 679 AVMAWAGPFSYLGIRSSLVILQVPLVSVCAADNIFIPVLEYQRLPRPGPREVHIGR 738
DB 316 GIASAMGMLTYLEIQYN-DIIAVMPFLVAVGTGDNMLVMSLAKRTDR--NLKYDQRIAE 372

QY 739 ALGRVAPSMCLLSSEACFFLGTALTPMPAVRTALTSLGLAVILDFLLQMSAFVALLSLD 798
Db 373 CMADAIVSILITDALSGFGVGTITTPAQVCIYTMCALLITFAQLTFFCAILVY 432
QY 799 SKROE-----ASRLDV---CCVKPQBLPPPGOG----- 824
Db 433 TRIEBQGLHSIWLRPVAVTSYTSPLNVKLFWLGSOPOK-PLPSCGTVSSTSVSTMTSQA 491
QY 825 -----EGLLGFFPKAYAPFLH-WITRGVLLILFLALFGVLSYSCMCHISVGLDOE 874
Db 492 TSPASKHLHCAATSFRRWYAPVLMQPMI-RATAGLWYLYLIGISYIGCTHLKEGLEPA 550
QY 875 LALPKDSYLLDYFELFNR-YPEVGPVYFVTT-----LGNFSEAGMNAICSSA- 923
Db 551 NLLVDDSYATPHRYVLEKHYHYGASIQIVVSNPPDLDPVERINMDKMASTFANCKVAI 610
QY 924 GGNFSPF-----TKIQYATE----FPEQSYLAIPASSWDDFDWLTPSSCCRL 969
Db 611 GDDSVQFWLREMQVSEETHKIQYDNEKFYDHAQAQYIYSDMQPWVUVV----- 659
QY 970 YISGENKOKFCPSTVNSLNCNKMCSITWGSVRPSVEQFHKYLPFLNDRPNKCPKGL 1029
Db 660 -----WGRNN----- 664
QY 1030 RAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPDTPA 1089
Db 665 -----NSRIKTRFMICMRDITTTKQTEATFRIASRF-----EQ 704
QY 1090 FEVFPYTTITNVEYQYLTLPGLFMLSCLVPTFAVSCLLGLDLRLSGLNLSIVMIL 1149
Db 705 YNVVTYMLPLTQDQYALVVPNTQDIIIVAVACMLVISALLIPQVCSFWV-AVTIGSID 763
QY 1150 VDTGFMALWDISYNAVSLINLVSAGMVSVEFVSHITRSPAISTKPTWLERAKEATISMG 1209
Db 764 LGVLGFMILMNVLDAISMITTIMSVGSFSDVYSNHTYAYVISKESITTSARVCDALGDLG 823
QY 1210 SAVFAGVAMTNLPGILVLGLAKAQIQIFFFFRLNLTLLGLLHGLVFLPVILS-YVG 1266
Db 824 NPVAQG-AMSTILAVSLSDVPAYMI-VTFFKTVFLAISIGLHGLVFLPLMLSVFVG 879
RESULT 75
Q4H343 CIOIN
ID Q4H343 CIOIN PRELIMINARY; PRT; 278 AA.
AC Q4H343;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Nienann-Pick disease, type C3 (Fragment).
GN Names-Ci-Niemann-Pick Type C3;
OS Clona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Cionidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15269171; DOI=10.1242/dev.01270;
RA Imai K.S., Hino K., Yagi K., Satoh N., Satou Y.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis."
RL Development 131:4047-4058(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=12736827; DOI=10.1007/s00427-003-0330-z;
RA Satou Y., Satoh N.;
RT "Genomewide surveys of developmentally relevant genes in Ciona
intestinalis."
RL Dev. Genes Evol. 213:211-212(2003).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RA Satou Y.;
RT "Expressed genes in Ciona intestinalis."
RN Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB210579; BAE06584.1; -, mRNA.
FT NON TER 1
SQ SEQUENCE 278 AA; 30653 MW; 59APAA65596844232 CRC64;
Query Match 8.5%; Score 589; DB 2; Length 278;
Best Local Similarity 47.2%; Pred. No. 3e-33;
Matches 120; Conservative 42; Mismatches 76; Indels 16; Gaps 6;
QY 869 VGLDQELALPKDSYLLDYFELFNRFEVGPVYFVTTLGNFSEAGMNAICSSAGCNF 928
Db 2 IGLDQSLMSEDSYLLDYFDMNNYLSVGAPVYFVVKDQGNQYTDAGANQICGGMGCNN 61
QY 929 SFTQKIQYATEFPEQSYLAIPASSWDDFDWLTP-SSCRLYISGNKDKFCPSTVNSL 987
Db 62 SLIQIARMSKMPNYSYIAVPASSWLDYFDWLKPQSSCCRHDTG-EEDVFCNATVVT 120
QY 988 NCLKNKMSI--TMGSVRPSVEQFHKYLPFLNDRPNKCPKGLAAYSTSVNLTSDG--- 1042
Db 121 SCIA-CRSAQESANQSRPTPDEPMKFLPFLNDRPNKCPKGLAAYSTSVNLTSDG 179
QY 1043 QVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPDTPAPEVEFPYTTINVFY 1102
Db 180 RVGATSMAYHTLTKTSKDFIGCLRSANKIAEIS-----QNTTAEVFPYVYVY 231
QY 1103 EQLTILPEGLFML 1116
Db 232 EQLTIVHDTIFNL 245
RESULT 76
Q61EN1 CAEBR
ID Q61EN1 CAEBR PRELIMINARY; PRT; 933 AA.
AC Q61EN1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein CBG12004.
GN Name=CBG12004;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC0100059; CAB66666.1; -, Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0007275; F:development; IEA.
DR InterPro; IPR001036; Acrlivin_res.
DR InterPro; IPR003392; Patched.
DR Pfam; PF02460; Patched; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
SQ SEQUENCE 933 AA; 104692 MW; 36364C782DC876DA CRC64;
Query Match 8.5%; Score 585; DB 2; Length 933;
Best Local Similarity 22.9%; Pred. No. 2.8e-32;
Matches 245; Conservative 186; Mismatches 370; Indels 268; Gaps 38;
QY 341 FQCGWGTWASWFLTILVLSVIPVVAALAGLVFTELTDPVLSAPNSQARSEK-AFH-- 397
Db 14 FRQLGFFVCDHPLFVFPVLLFTTAMGVLLHNLPLSDAVLFTPTGAQSKMERMSIHEK 73
QY 398 ----DOHFGP----PFRNQVILTAPNRSRYRVSLLGPKNFSGILDLLELLELQER 450

[MIM:155255]. MDB is a malignant, invasive embryonal tumor of the cerebellum with a preferential manifestation in children. Although the majority of medulloblastomas occur sporadically, some manifest within familial cancer syndromes such as Turcot syndrome and basal cell nevus syndrome (Gorlin syndrome).

-1- DISEASE: Defects in PTCH2 are a cause of sporadic basal cell carcinoma (BCC) [MIM:605462].

-1- SIMILARITY: Belongs to the patched family.

-1- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.

This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.

EMBL; AF091501; AAC79847.1; -; mRNA.

EMBL; AF087651; AAD25953.1; -; mRNA.

EMBL; AF119569; AAD17260.1; -; mRNA.

EMBL; AY385555; AAQ88919.1; -; mRNA.

EMBL; AY438664; AAR05447.1; -; Genomic DNA.

EMBL; AL136380; CAI23127.1; -; Genomic DNA.

EMBL; AL592166; CAI23127.1; JOINED; Genomic DNA.

EMBL; AL592166; CAI13000.1; -; Genomic DNA.

EMBL; AL136380; CAI13000.1; JOINED; Genomic DNA.

Ensembl; ENSG00000117425; Homo sapiens.

DR HGNC; HGNC:9586; PTCH2.

DR MIM; 603673; -.

DR MIM; 155255; -.

DR MIM; 605462; -.

DR GO; GO:0005887; C:integral to plasma membrane; TAS.

DR GO; GO:0004886; F:transmembrane receptor activity; TAS.

DR GO; GO:0008544; P:epidermis development; TAS.

DR GO; GO:0006461; P:protein complex assembly; TAS.

DR GO; GO:0007283; P:spermatogenesis; TAS.

DR InterPro; IPR003392; Patched.

DR InterPro; IPR004766; Patchedtm_recept.

DR InterPro; IPR000731; SSD_5TM.

DR Pfam; PF02460; Patched; I.

DR TIGRFAMs; TIGR00918; 2A060602; 1.

DR PROSITE; PS50156; SSD; 1.

KW Alternative splicing; Glycoprotein; Polymorphism; Receptor;

KW Transmembrane.

FT TOPO_DOM 1 57 Cytoplasmic (Potential).

FT TRANSMEM 58 78 Potential.

FT TOPO_DOM 79 392 Extracellular (Potential).

FT TRANSMEM 393 413 Potential.

FT TOPO_DOM 414 428 Cytoplasmic (Potential).

FT TRANSMEM 429 449 Potential.

FT TOPO_DOM 450 457 Extracellular (Potential).

FT TRANSMEM 458 478 Potential.

FT TOPO_DOM 479 501 Cytoplasmic (Potential).

FT TRANSMEM 502 522 Potential.

FT TOPO_DOM 523 531 Extracellular (Potential).

FT TRANSMEM 532 552 Potential.

FT TOPO_DOM 553 586 Cytoplasmic (Potential).

FT TRANSMEM 587 707 Potential.

FT TOPO_DOM 708 963 Extracellular (Potential).

FT TRANSMEM 964 984 Potential.

FT TOPO_DOM 985 991 Cytoplasmic (Potential).

FT TRANSMEM 992 1012 Potential.

FT TOPO_DOM 1013 1013 Extracellular (Potential).

FT TRANSMEM 1014 1034 Potential.

FT TOPO_DOM 1035 1064 Cytoplasmic (Potential).

FT TRANSMEM 1065 1085 Potential.

FT TOPO_DOM 1086 1093 Extracellular (Potential).

FT TRANSMEM 1094 1114 Potential.

FT TOPO_DOM 1115 1203 Cytoplasmic (Potential).

FT DOMAIN 394 552 SSD.

FT CARBOHYD 370 370 N-linked (GlcNAc...) (Potential).

FT CARBOHYD 812 812 N-linked (GlcNAc...) (Potential).

FT VARSPPLIC 1143 1203 WGASSSLPQSFAVVTSMTVAIHPPGLGAYTHPAPDEPPFW SPAAATSGNLSRGGPGPATG -> PEEI (in isoform

FT 346 346 /FTid=VSP_004542.

FT VARIANT E -> Q (in dbSNP:11573578).

FT 493 493 /FTid=VAR_018935

FT VARIANT E -> K (in dbSNP:11573581).

FT 622 622 /FTid=VAR_018936.

FT VARIANT H -> Y (in dbSNP:11573586).

FT 988 988 /FTid=VAR_018937.

FT VARIANT T -> M (in dbSNP:11573590).

FT 1019 1019 /FTid=VAR_018938.

FT VARIANT V -> M (in dbSNP:11573591).

FT 1121 1121 /FTid=VAR_018939.

FT VARIANT I -> M (in dbSNP:11573598).

FT 2 2 /FTid=VAR_018940.

FT CONFLICT T -> N (in Ref. 2).

FT 120 120 R -> L (in Ref. 2).

FT CONFLICT 175 175 R -> W (in Ref. 1 and 4).

FT CONFLICT 753 753 Q -> R (in Ref. 2).

FT 787 787 P -> S (in Ref. 2).

FT CONFLICT 833 833 P -> L (in Ref. 3).

FT 837 837 S -> G (in Ref. 2).

FT CONFLICT 846 846 D -> E (in Ref. 2).

FT 897 897 L -> F (in Ref. 3).

FT CONFLICT 1203 1203 G -> GDYKDDDK (in Ref. 4).

SQ SEQUENCE 1203 AA; 130544 MW; 4FAB06999782C031 CRC64;

Query Match 8.5%; Score 584; DB 1; Length 1203;

Best Local Similarity 22.9%; Pred. No. 4.7e-32;

Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;

QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGQFFQG---WGTWVASWPLTILVLS 359

DB 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLRAVPGQLLFSLGGCIORHCGKVLFG 62

QY 360 VIPVVALAAGLVFTLTDVPELMSAPNSQARSEKAFHDHFG--PFRTNQVILTAPNR 417

DB 63 LLAFGALGALGRMAIIEIETNLEQLWVEGVSQELHYTKELKGBEAAVTSQMLQTARQE 122

QY 418 SSVRYDLSLLGPKNFSGILDLLELELERLHLQVWSPQARNISLQICV---AP 474

DB 123 GEN-----ILTPALG-----LHQAALTASKVQVSLYKGSWLNKIKYKSGVP 166

QY 475 LNPDN-----TSLYDCCINSLQYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520

DB 167 LIENGMIERMTEKLPCCVILTPDLCFMEGAK---LQGGSAYLPGRPDIQWNLDPQLLE 223

QY 521 -----HFLYCANAPLTPKDGITALALSCMADYGA 548

DB 224 ELGPPASLEGPRELLDKAQVGQAVYVGRPCILHPDDDLHCPSPAPNHHSRQAPNVVHLSGGC 283

QY 549 PVFPP-----LAIGYKGYKDYSE---AEALIMTF-----SLNNTYPAGDPRLAQAKL 591

DB 284 HGFSKHFHWQBEILLGGWARDPQGEELRAELAQSTFLMSPRQLYEHFRGDTYQTHDIGW 343

QY 592 WEEAFLEEMRAPQRMAGMFQVTFARSLR-DEINRTAEDLPFATS-----Y 640

DB 344 SEEQASTVLQAWQRRFVOLAQELPENASQQTHAFSSTLTDDI-LHAFSEVSAAARVWGY 402

QY 641 IVIFLYISLALGYSYSSRWVVDSKATLGLGVAVVLGAVMAAGFFSVLGRSSRLVLIQ 700

DB 403 LLMLAYACVTM---LRWD---CAQSGSVGLAGVLIVLAVASGLGLCALLGITFNATQ 457

QY 701 VVPFLVLSVGADNIFVLEYQRLPRRPFGEPEVHIGRALGRVAPSMMLCSISEAICFFPL 760

DB 458 VLPFLALGIGVDDVFL--LAHAFTEALPGTFLQERMGECQLQTGTSVLTSLNNMAAFLM 515

QY 761 GALTWPVAVRTPALTSGLAVIDLFLQMSAFVALLSLDSKQESRLDYCCC----- 812

DB 516 AALVIPALRAFSLOAAIIVWGCTFVAVMLVFPAILSLDLRRRCORLDVLCFSSPCSAQ 575

QY 813 ----VKPQEL-----PPQCGSEGL-- 827

DB 576 VIQILPQELGDDGTVPGVGINHLTATVQAFTHCEASSQHVVTLPPQAHVLPFPPSDPLGSEL 635

QY 828 -----LLG-----FFKAYAPFLLHMITRGVLLFLA 855
 Db 636 FSPGSTRDLLOQBETROKAAKSLPCARWNLAHFARYQFAPLQLQSHAKAIVLVFGA 695
 QY 856 LFGVSLYSNCHISVGLDQELALPKDSYLLDYFLNRYFEVGCAPYFVTTILGYNFS-SEA 914
 Db 696 LGLSLYGATLVQDGLATDVVPRGTKEHAFSLAQLRYFSL-YEVALVTQGGFDVAHSQR 754
 QY 915 GNAICSSAGCNFSFTQKIQYATFPEQSYLAIPAS-----SWDDFDIMLTPSSCC--- 967
 Db 755 AL-----FDLHQRFSSL-----KAVLPPEATQAPRTWLHYRNWLOGIOQAADQ 798
 QY 968 -----RLYISGPNKDKFCPSTVNSLNCNKMSTMGSVRSVQFHKYLPWFNLD 1018
 Db 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQBPLD-FSQTTRKLVLD 846
 QY 1019 RNIKCPKGGLAAYSTVNLSDGOVLASRFMAYHKKP----- 1055
 Db 847 REGLIPPE--LFYMGLTWVSSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQ 904
 QY 1056 -----LKNQSDYTEALRAARELAANI-TADLRKVPGTDPAFEVFPYTTINVFY 1102
 Db 905 LEPAQFPFLRLGLQKTADFVEAIEGARAACAEAGQGVHAYPSGSPF-----LFW 954
 QY 1103 EQLYTLPEGLPMLSLC--LVPTFAVSCILLGLDLRSLGLNLSIVMLVDTVGFPMALWD 1160
 Db 955 EQLYGL--RRCPFLAVCILLVCTFLVALLNPNWTAGLI-VLVAMTVELFGIMGFLG 1011
 QY 1161 ISYNAVLINLVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
 Db 1012 IKLSAIPVILVASGVIGEFVHVVALGF-LTQGSRLNRAHAHALEHTFAPVTDG-AIST 1069
 QY 1221 LPIGLVLGAKAQLIQIPFFRLNLLITLLGLLHGLVFLPVILSVYVGPDPNPALALEQKRA 1280
 Db 1070 LLGLMLAGSHDFIVRFFAALTVTLLGLLHGLVLLPVLSILGPP--PEVIMYKES 1127
 QY 1281 EBAVAAV-----MVASCNPHSRVSTADNIYVN 1308
 Db 1128 PEILSPPAQGGGLRWGASSLSLPQSFARVTTSMVTVAIH 1165
 RESULT 78
 ID Q53257 HUMAN PRELIMINARY; PRT; 1203 AA.
 AC Q53257
 DT 13-SEP-2005 (TrEMBLrel. 31, Created)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
 DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
 DE Patched 2.
 GN ORFNames=UNQ560;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D.T., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J.S., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klinowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M.R., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R.L., Watanabe C., Wiedand D., Woods K., Xie M.-H.,
 RA Wood W.D., Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A.D.,
 RA Wood I., Godowski P.J., Gray A.M.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment."
 RL Genome Res. 13:2265-2270(2003).
 DR EMBL; AY359016; AAQ89375.1; -; mRNA.

SQ SEQUENCE 1203 AA; 130573 MW; FFFA069997EEB6EB CRC64;
 Query Match 8.5%; Score 584; DB 2; Length 1203;
 Best Local Similarity 22.9%; Pred. No. 4.7e-32;
 Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;
 QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGQFFQG-----WGTWVASHPLFILVLS 359
 Db 3 RSPPLRELPPSTTPARTAAQILAGSLKAPLWRAFYQGLFLSLGCGIQRHCGKVLFLG 62
 QY 360 VIPVVALAAGLVFTETLTPDVELSNAPNSQARSEKAFHDQHEG--PFFRTNQVILTAPNR 417
 Db 63 LLAFGALGALGRLMAIETNLNQLWVEGSRVSQELHYTKELGEBAAVYSQMLIQTAROE 122
 QY 418 SSYRDSILLGPKNPSGILDLDLLELELERLHLQVMSPEAQRNLSLQDICY---AP 474
 Db 123 GEN-----LHQAALTASKVQVSLYKGSWDLNKICYKSGVP 166
 QY 475 LNPN-----TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520
 Db 167 LIENGMIEMIEKLPFCVILTPDLCFWECAK---LQGGSAVLPGRPDIQWTLNLDPEQLIE 223
 QY 521 -----HELYCANAPLTFKDGTAALSLSCMADYCA 548
 Db 224 ELGPPASLEGFRELDKAQGVYVGRPCLLHPDDLHCPSPNHHRSQAPNVHLSGGC 283
 QY 549 PVFPF-----LAIGYKGYKDYSE---ABALIMTF-----SLNNTYPAGDPRLAQAKL 591
 Db 284 HGFHKFPMHQBELLLGGWARDPOCELLRAELQSTFLLMSPRQLYEHFRGDYQTHDIGW 343
 QY 592 WEBALEEMRAPFORMAGMFQVTTFAERSLE-DEINRTTAEDLPPIFATS-----Y 640
 Db 344 SEQAQSTVLQAWRRFVQLAQLPENASQIHAFTSSTLDDI-LHAFSEVSAARVWVG 402
 QY 641 IVIFLYISLALGSSVSSWRVVDKATGLGCVAVVLGAVMAAMGFFSVGLTRSSILVILQ 700
 Db 403 LMLLAYACVTM---LRWD--CAQSGSVGLAGVLLVALAVASGGLCALLGTFAAATQ 457
 QY 701 VVPLVLSVAGADNIFIVLEYORLPRRPGEPREHVHIGRALGRVAPSMILCSLEAICPPL 760
 Db 458 VLPFLALGIGVDDVPL--LAHFTALPGTFLQERMGECLOQTGTSVVLTSINNMAFLM 515
 QY 761 GALTMPAARTFALTSLGLAVILDFLQMSAFVALLSLDSKROESRLDVCC----- 812
 Db 516 AALVPIPALRAFSLQAAIVVGCTFVAVMLVFPAILSLDLRRHCQLDLVLCFSSPSCSAQ 575
 QY 813 ---VKPQEL-----PPGQGEGL-- 827
 Db 576 VIQILPQELGDTVPVGIHILTATVQAFTHCEASSQHVVTILPPOAHLVPPSPDPLGSEL 635
 QY 828 -----LLG-----FFOKAVAPPLLHMITRGVLLFLA 855
 Db 636 FSPGSTRDLLOQBETROKAAKSLPCARWNLAHFARYQFAPLQLQSHAKAIVLVFGA 695
 QY 856 LFGVSLYSNCHISVGLDQELALPKDSYLLDYFLNRYFEVGCAPYFVTTILGYNFS-SEA 914
 Db 696 LGLSLYGATLVQDGLATDVVPRGTKEHAFSLAQLRYFSL-YEVALVTQGGFDVAHSQR 754
 QY 915 GNAICSSAGCNFSFTQKIQYATFPEQSYLAIPAS-----SWDDFDIMLTPSSCC--- 967
 Db 755 AL-----FDLHQRFSSL-----KAVLPPEATQAPRTWLHYRNWLOGIOQAADQ 798
 QY 968 -----RLYISGPNKDKFCPSTVNSLNCNKMSTMGSVRSVQFHKYLPWFNLD 1018
 Db 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQBPLD-FSQTTRKLVLD 846
 QY 1019 RNIKCPKGGLAAYSTVNLSDGOVLASRFMAYHKKP----- 1055
 Db 847 REGLIPPE--LFYMGLTWVSSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQ 904
 QY 1056 -----LKNQSDYTEALRAARELAANI-TADLRKVPGTDPAFEVFPYTTINVFY 1102
 Db 905 LEPAQFPFLRLGLQKTADFVEAIEGARAACAEAGQGVHAYPSGSPF-----LFW 954

QY 1103 BOYLTLPEGLFMLSCL--LVPTAVSCLLLGLDLRSLGLNLSTVMLVDTVGFMAWD 1160
 Db 955 BOYLGL--RRCLLAVCLLLVCTFLVCAALLNFWTAGLI-VVLAMMTVSLFGLMGFLG 1011
 QY 1161 ISYNAVSLINVSAGMSEVFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
 Db 1012 IKLSAIPVVLVASGIGVEFTVHVALGF-LTQGSNLRRAHALEHTFAPVTDG-AIST 1069
 QY 1221 LPGAIVLGLAKAQLIQIPFFRLNLITLGLLHGLVFLPVLVSYPVDPVNPALALEQKRA 1280
 Db 1070 LGLMLAGSHDFDFTVRYFFAALTVLTLGLLHGLVFLPVLVSILGPP--DEVIQMYKES 1127
 QY 1281 EENAAV-----NVASCPNHPRSVSTADNIYVN 1308
 Db 1128 PEILSPAPQGGGLRWGASSLSPOSFAFVTTSMTVAIH 1165

RESULT 79
 PTC2_MOUSE STANDARD; PRT; 1182 AA.
 AC O35595; Q54670;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Patched protein homolog 2 (PTC2).
 GN Name=Ptch2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridea; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA MEDLINE=98122566; PubMed=9462734;
 RA Motoyama J., Takabatake T., Takekuma K., Hui C.-C.;
 RT "Ptch2, a second mouse Patched gene is co-expressed with Sonic
 hedgehog.";
 RL Nat. Genet. 18:104-106(1998).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [GENOMIC DNA].
 RC STRAIN=129/SvJ; TISSUE=Liver;
 RX PubMed=12438747; DOI=10.1159/000064064;
 RA Froehlich L., Zhanqu L., Beier D.R., Lanke B.;
 RT "Genomic structure and refined chromosomal localization of the mouse
 Ptch2 gene.";
 RL Cytogenet. Genome Res. 97:106-110(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE [MRNA] OF 196-446.
 RC STRAIN=BALB/c; TISSUE=Neuroretina;
 RX MEDLINE=97379366; PubMed=9237688; DOI=10.1016/S0014-5793(97)00645-5;
 RA Takabatake T., Ogawa M., Takahashi T.C., Mizuno M., Okamoto M.,
 RA Takeshima K.;
 RT Hedgehog and patched gene expression in adult ocular tissues.";
 RL FEBS Lett. 410:485-489(1997).
 CC -!- FUNCTION: May have a role in epidermal development. May act as a
 receptor for Sonic hedgehog (SHH).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in epithelial cells of the
 developing hair, tooth and whisker.
 CC -!- DEVELOPMENTAL STAGE: Detected in 8.5 to 17.5 dpc embryos.
 CC -!- SIMILARITY: Belongs to the patched family.
 CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.
 CC
 CC EMBL; AJ133482; CAC88120.1; -; Genomic DNA.
 CC EMBL; AJ133483; CAC88120.1; JOINED; Genomic_DNA.

DR EMBL; AJ133484; CAC88120.1; JOINED; Genomic DNA.
 DR EMBL; AJ133485; CAC88120.1; JOINED; Genomic_DNA.
 DR EMBL; AB010833; BAA24691.1; -; mRNA.
 DR PIR; I13952; I13952.
 DR Ensembl; ENSMUSG00000028681; Mus musculus.
 DR MGI; MGI:1095405; Ptch2.
 DR GO; GO:0016021; C:integral to membrane; TAS.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR004766; Patchedtm_recept.
 DR InterPro; IPR000731; SSD_SFM.
 DR Pfam; PF02460; Patched; 1.
 DR TIGRFAMs; TIGR00918; 2A060602; 1.
 DR PROSITE; PS0156; SSD; 1.
 KW Glycoprotein; Receptor; Transmembrane.
 FT TOPO_DOM 1 57 Cytoplasmic (Potential).
 FT TRANSMEM 58 78 Potential.
 FT TOPO_DOM 79 394 Extracellular (Potential).
 FT TRANSMEM 395 414 Potential.
 FT TOPO_DOM 415 428 Cytoplasmic (Potential).
 FT TRANSMEM 429 449 Potential.
 FT TOPO_DOM 450 457 Extracellular (Potential).
 FT TRANSMEM 458 478 Potential.
 FT TOPO_DOM 479 501 Cytoplasmic (Potential).
 FT TRANSMEM 502 522 Potential.
 FT TOPO_DOM 523 531 Extracellular (Potential).
 FT TRANSMEM 532 552 Potential.
 FT TOPO_DOM 553 586 Cytoplasmic (Potential).
 FT TRANSMEM 587 707 Potential.
 FT TOPO_DOM 708 963 Extracellular (Potential).
 FT TRANSMEM 964 984 Potential.
 FT TOPO_DOM 985 991 Cytoplasmic (Potential).
 FT TRANSMEM 992 1012 Potential.
 FT TOPO_DOM 1013 1013 Extracellular (Potential).
 FT TRANSMEM 1014 1034 Potential.
 FT TOPO_DOM 1035 1064 Cytoplasmic (Potential).
 FT TRANSMEM 1065 1085 Potential.
 FT TOPO_DOM 1086 1086 Extracellular (Potential).
 FT TRANSMEM 1087 1107 Potential.
 FT TOPO_DOM 1108 1182 Cytoplasmic (Potential).
 FT DOMAIN 394 552 SSD.
 FT CARBOHYD 370 370 N-linked (GlcNAc..) (Potential).
 FT CARBOHYD 812 812 N-linked (GlcNAc..) (Potential).
 SQ SEQUENCE 1182 AA; 128586 MW; 715233D912C352F2 CRC64;

Query Match Similarity 8.4%; Score 593.5; DB 1; Length 1182;
 Best Local Similarity 24.2%; Pred. No. 5e-32;
 Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SFSTHTLLGQ-----PFQG-----WGTVVASWPLTILVLSVPPVVALAAGLVFTLT 376
 Db 20 SSAPHILAGSLQAPLWLRAYFQGLLFLSGCRQKHCGKVLFLGLVAFGLALGLRVAVIE 79
 QY 377 TDPVELMSAPNSQARSEKAFHDQHFQ-PFFRTNQVILTPAPNRSSVRYDLSLLGPKNFSGI 435
 Db 80 TDLEQLWVEGSRVSRVQELHYTKELGEEAAVTSQMLI-----QTAHQEGGNVLTP-----A 131
 QY 436 LDLDLLELLELQERLRLHQLWSPQAQRNLSIDICY---APLNPDN-----TSLYDCC 486
 Db 132 LD-----LHQAALTASKQVSLYSGKSWDLNKLICYKSGVPLIENGMIERMIKLPFCV 184
 QY 487 INSLLOYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---FKDGTALA 539
 Db 195 ILTPLDCEWGAQK---LOGGSAYLPGRPDIQNTWLDPOQLLEELGPFASLEGFRELLDKA 241
 QY 540 LSCMDADYGAPVF-----PFLA---IGGYKGDYS----- 565
 Db 242 QVGQAYVGRPCLDPDHPCHPSPAPNRHSRQAPNVAQELSGGCHGFSHKFMHWQEEILLGG 301
 QY 566 -----BAEALIMTF-----SLNNYPAGDPRLAQAKLWEEAFLEEMAFORFWMAG 609
 Db 302 TARDLQGLLRALQSTFLLMSRPLQYEHFRFGDQTHDIGWSESEQASWQAWQRRFVQ 361
 QY 610 MFQVTFTAERSLE-----DEINRTTAEDLPFATSYIV--IFLYISLALGSYSWS 658

```

Db 362 LAQALPANASQIHAFTSTTLDILRAFSE-----VSTTRVGGYLLMLAYACVTMLRWD 417
QY 659 RYMVDSKATLGLGGVAVVLGAVMAAGFFSYLGRSSVLQVPPFLVLSVGADNIFIV 718
Db 418 --CAQSQAGVLAGVLLVALAVASGLICALLGITFNAATTQVLPPLALGIGVDDIFLLA 475
QY 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFIGALTPMPAVRTFALTSL 778
Db 476 HAFKAP--PDTPLPERMGECLRTGTSTGVALTSVNNVAFVMAALVPIPALRAFSIQAI 533
QY 779 AVILDFLLQMSAFVALLSDSKOERSLDVCC-----VKPOE-----817
Db 534 VVGCNFAAVMLVFPALLSLDLRRHRQRDLVLCFSPSPCSAQVQKLPQELGRDRAVPVGI 593
QY 818 -----LPP-----PG-----826
Db 594 AHLTATVQAFTHCEASSQHVVLFPQAHLLSPASDPLGSELSPGSTRDILLSQEGTG 653
QY 827 -----LLGPFQKAYAPPLLHWITRGVLLFLALFGVLSYMSCHISVGLDQ 873
Db 654 PQACRPLLCALHTLAHFARYQFAPLLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713
QY 874 ELALPKDSYLLDYFLFLNRYEVGAPVYVTTLYNFS--SEAGMAICSSACNPFSTQ 932
Db 714 TDVPRGTKEHAFLSAQLRYFSL--YEVALVTGGFDYAHQSRL-----FDLHQ 761
QY 933 KIQYATEPPEQSYLAIPAS-----SWVDDF-----IDWLTPSSCCRLYISGPNK 976
Db 762 RFSSL-----KAVLPPATQAPRTWLHYRSMLOGIAQAFDQDQWASGRITCHSYNGSED 816
QY 977 D-----KPCSTVNSLNCNKMSTW-----GSRVPSVEQPHKYPFLNDRPNKCPK 1026
Db 817 GALAYKLAIQTGNAQEPDFSLQTRKLVKDEGLIPP--ELFYMGLTVMVSSDPL-----869
QY 1027 GGLAA-----YST-----SVNLTSQGVLASRFMAHYKPLKNSQDTEALR 1067
Db 870 --GLAASQANFYPPPEWLHDKYDTTGENLRIPAAQPLEPAQFPFLHLGLQKTDADFEAIB 928
QY 1068 AARELAANI--TADLRKVPCTDPAFEVFPVTITNVFEQVLTLPBGLFMLSIC--LVPTF 1124
Db 929 GARAACTEAGQGVHAYSGSPF-----LFEQYGLG--RRCFLAVCILLVCTF 976
QY 1125 AVSCLLGLDRLSGLLNLISVIMLVDTVGFMAWDISTYNSVLINLVSAGMSYFVSH 1184
Db 977 LVCALLLSPTWAGLI--VLVAMMTVELFGIMFLGIXLSAIPVVLVASIGIGVEFTVH 1035
QY 1185 ITRSPAISTKPTWLERAKEATISMSGSAVPAGV---AMTNLPGLVILGLAKAQLIQIPFR 1241
Db 1036 VALGSLTSHGSRNRAA-----SALEQTPAPVTDGAVSTLLGLMLAGSNFDFIIRYFV 1090
QY 1242 LNLITLTLGLLGLVFLPVILSYVGP 1267
Db 1091 VLTVLTLLGLLGLLPLVLLSILGP 1116

RESULT 80
Q9XXR9 CAEL
ID Q9XXR9_CAEL PRELIMINARY; PRT: 1003 AA.
AC Q9XXR9;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Hypothetical protein ptr-19.
GN Name=ptr-19; ORFNames=Y39A1B.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RX [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;

```

```

RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: AL021482; CAA16339.1; -; Genomic_DNA.
DR PIR: T26746; T26746.
DR Ensembl: Y39A1B.2; Caenorhabditis elegans.
DR WormBase; WBGene0004233; Y39A1B.2.
DR WormPep; Y39A1B.2; C19144.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1003 AA; 112257 MW; 92EAP1EA1P7627C CRC64;

Query Match 8.3%; Score 573.5; DB 2; Length 1003;
Best Local Similarity 22.8%; Pred. No. 2e-31;
Matches 236; Conservative 198; Mismatches 376; Indels 225; Gaps 42;

QY 339 OFFQCGTW---VASWPLTILVLSVIPVVALAAGLVFTE--LTTDPVELWSAPNSQARSE 393
Db 9 RPAHAFGAYGVVVRHPPFPFPIIPILITAGLSTGLRHEQAFMKDELELYTPTDAQKE 68
QY 394 KA-----FHDQHFQGFRTNQ-----VILTAENRSYRYDLSLLGKPNFSGILDLDL 440
Db 69 LSQDLHLFHINDSDFYATRRYDIRRAGYIIVT-----NOEDGDIINPLVMSAQLWS 123
QY 441 LLELELERLHLQVMSPEAQRNLSQDIC-----YAPLNPNTSYDDCCI 487
Db 124 IVQSLTVED-----EDRRINYPSCIVKPFIPPEFSKALHSLFAPNMTTPEECV 173
QY 488 -NSLQYQNNRTLLLTANQTLMGQT-----SQVMDKHLYCANAPLTKDGTALAL 540
Db 174 SNPLVEIFK-----LLVSDRSFLNRSIDEMTLSQI-----SDAIQFDSGGMTHL 218
QY 541 SCMDYGAVPFPFLAIGY---KGQDYSEAEALIMTFSLLNYPNYPAGDPRLAQAKLWEEAFL 597
Db 219 -----LGGVTLDDDDKRIAGAKAMLLPYALRH--SSDDEWDVAEKVEVRLA 261
QY 598 BEMAFQRMAGMFOVFTFAERSLDEINRTTAPDLPIFATSYIVIFLYISLALSGYSW 657
Db 262 DFLQYD---SPIIRASWMTYETLAAESARDRLQIHMLLPCFVCVSIPTIACCCVFSW 317
QY 658 SRVWDSKATLGLGVAVVLGAVMAAGFFSVLGRSSVLQVPPFLVLSVGADNIFIF 717
Db 318 RR-----SRPLAIGGVISAAMAIASAVGILLAGYGMTSVAYS--MPFIVFSYGVNDVFL 372
QY 718 VLEYORLPRRPGPREVHIGRALGRVAPSMLLCSLSAICFFIGALTPMPAVRTFALTSG 777
Db 373 LSAWRSTST--ETLEHRMKETFPADAIVSTLTDLISFGVGCATFPFVSQMCAYAV 430
QY 778 LAVILDFLLQMSAFVALLSDSKOERSLDVCCV-----KPQELPPPGQEG-----826
Db 431 AAVIFTYIYQLTFPAAVVMTNRREINR-----HCIPPHKLKQDTLPEKIAAQGRSPK 486
QY 827 --LLGPFQKAYAPPLLHWITRGVLLFLALFGVLSYMSCHISVGLQELALPKDSYLL 884
Db 487 NTILAQFFRTTYSDFLLNPLVRIVLTVFCVVLGVASYGCTKVKLGLSPNDLLPNSYK 546
QY 885 DYFLPLNRYF--EVGAPVYVTTLYNFS--BAGMNAICSSAGCNFSTFKIQIYATERPE 942
Db 547 RTLMAEKYFSDYSGSLH---VWMTNLSVDVAPRKIWN-----VLEKEVELYE 592
QY 943 QSYLAIPASSWVDDFDW-----LTPSSCCRLYISGPNKDKFCPTVNSLNCNKM 995
Db 593 HTEFTASSDSMLRTPLAFVKQAGLLITPE-----NFVYLKN---629
QY 996 ITMSVRSVPSTQPHKYPFLNDRPNKCPKGLLAAYSTSVNLTSQGVLT-ASRFMAHYK 1054
Db 630 -----VFLSQPF-----AKYNRDVVLTEDGBHLEASRI-----658

```



```

Db 444 D---CAKSGAGVAGVLLVALSVAGLGLCSLLGLSFFNAATQVLPSTLALGIVDDMF- 500
QY 718 VLEYQRLPRPPEVHIGRALGRVAPSMMLCSLSAICFFLIGALTPMPAVRTFALTSG 777
Db 501 -LGHSTETSRNPCKERTGDCRLRTGTSTVALTSVNNMIAFFMAALVPFIPALRAFSLQRA 559
QY 778 LAVILDFLLQMAFAVALLSLDSKQASRLDVCCC-----VKQOE----- 817
Db 560 VVVVFNFAMALLIFAILSLDLHREDKRLDILCCFPYSPCSRVIQIQOELSDANDNHQ 619
QY 818 -----LPPPGQ----- 823
Db 620 RAPATPTTGTSTTTTTHTTVQAPTQDAAAGQHIVTILPSTQSLSTTPPSSMWLSTPTP 679
QY 824 -----GEGE-----LLGFFQKAYAPFLHWTIR 846
Db 680 TTDYGSQVPTSSSTRDLAQVERPKEGRECVPPLPFFRNWSSFAREKAYALLKPKTK 739
QY 847 GVVLLEFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFVGVAPVYFVTTL 906
Db 740 TVVVVVFALLSLSLYGTMTWHDGLYLTDIVPRDTQEVYFITAQPKYFSF-YNNYLVMTD 798
QY 907 GYNFSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLA-----IPASSWVDVDFIDW 960
Db 799 GFDIAR-----SORQLQLHNAFNSVKYVVDKGNHKLK-P-RMWLHYFQDW 841
QY 961 LTPSSCC-----RLYISGPNKDKPCPSTVNSLNCCKNC 993
Db 842 LKGLQATFDADWEAGKITVDSYRNGTGDGALAYKPLIQTSKKPEPNYSQLSRRLVDG- 900
QY 994 MSITGWSVRPSVEQHFHYLP-WFLNDRPNKCPKGGALAA-----YST----- 1034
Db 901 -----DGLIPEV--FYIYLTWVSND-----PLGYAASQANFYPHPREMIDKYDTTGE 948
QY 1035 SVNLTSQGVLASRPMAYHKPKLNSQDYTEALRAARELAN-ITADLRKVPCTDPAFVFP 1093
Db 949 NLRIPAAEPLBPAQFPFYLNGLRQASDFTEATESVRTICEEPMQCIKNYPNG-----Y 1002
QY 1094 PYTITNVFYEQVLTILPEGLFWLSCLVPTPAVSCLLGLDLSRLSGLLNLSIVMLVADTV 1153
Db 1003 PF-----LFEQYIGURHWFLLSISVVLAFTVCAILLNPNWAGVI-VFILPMVTVELF 1057
QY 1154 GPMALWDISYNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSA 1211
Db 1058 GIMGLIGIKLSAIPVILLIASVIGVEFTVHALGFLTAIGDNT--RSAVAMEHPAP 1114
QY 1212 VPAGVAMTNLPGLVLGLAKAQIQLIFFRLNLLITLGLLHGLVFLPLVILSYVGP--DV 1269
Db 1115 VIDG-AISTLLGVLMAGSEFDFIMRYFPFAVLAILTLGLILNGLVLLPVLLSLMGPPABV 1173
QY 1270 NEA 1272
Db 1174 VFA 1176

```

RESULT 82

Q60YB2_CAEBR

ID Q60YB2_CAEBR PRELIMINARY;

AC Q60YB2; PRT; 1004 AA.

DT 25-OCT-2004 (TrEMBLrel. 28, Created)

DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)

DE Hypothetical protein CBGI18310.

GN Name=CBGI18310;

OS Caenorhabditis briggsae.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6238;

RN (1)

RP NUCLEOTIDE SEQUENCE.

RG The C.briggsae Sequencing Consortium;

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAAC01000088; CAB71402.1; -; Genomic DNA.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0008158; P:hedgehog receptor activity; IEA.

DR GO; GO:0007275; P:development; IEA.

DR InterPro; IPR0003392; Patched.

DR InterPro; IPR000731; SSD 5TM.

DR Pfam; PF02460; Patched; 1.

DR PROSITE; PS50156; SSD; 1.

DR Hypothetical protein.

KW SEQUENCE 1004 AA; 112493 MW; 4759FBI9FBSLIA6 CRC64;

Query Match 8.2%; Score 566.5; DB 2; Length 1004;

Best Local Similarity 22.8%; Pred. No. 6.4e-31;

Matches 237; Conservative 198; Mismatches 372; Indels 231; Gaps 43;

QY 339 QFFQGWGW---VASWPLTILVLPVVVALAAGLVFTS--LTTDPVELMSAPNSQARSE 393

Db 9 REAHAFGAYSGVVVRHPPFFIVPILITAAALTGLFRHEQAQFWKDELELYTPTDAQARKE 68

QY 394 -----KAFHQHGFPPFTNQ-----VILTAPNRSSRYRDSLLGLPKNFSGLDLDL 440

Db 69 LDQLDLRFLHINDSDPFYATRRYDIRRAGYIIVT-----NOEDGDILNPLVMHAAMQLWS 123

QY 441 LLELELEQLERHLQVWSPQAORNTSLQDIC-----YAPLNPDNTSLYDCCI 487

Db 124 IVQSUTVED-----EDRRINYPISCVKPIPPPEFSKALHSLFSPNMTTPEICV 173

QY 488 -NSLLQYFQNNRTLLLTANQTLMGQT-----SQVDMKDFLYCANAPLTKDGTALAL 540

Db 174 SNPLVEIFEK-----LLLVSDRSFLNRSIDEMTLSQI-----SDAIQPDGSGMTHL 218

QY 541 SCADYGAVPFPFLAIGGY---KGDKYSEAEALIMFTSLNNYPAGDPRLAQAOKLWEAEFL 597

Db 219 -----LGGVTLDLDDKRIAGAKAMLLFYALRH--SSDDEDDVAAEKWEVRLA 261

QY 598 BEMRAFQRMAGMFQVTTAERSLDEINRTTAEOLPIFATSVIVIFLYISLALGSSWS 657

Db 262 DFLQLYD---SPILIRASMTWTETLAAESARDRLQINMLLPCFVCVSIY-TIACCVCVFSW 317

QY 658 SRVMVDSKATLGLGVAVVLGAVMAAMGFPSYLGIRSSILVILQVVPFLVLSGADNIFIF 717

Db 318 RR-----SRPLAIGGVISAAMAIASAVGLLLAGYGMTSVAYS-MPFIYFVSVDNVFIL 372

QY 718 VLEYQRLPRPPEVHIGRALGRVAPSMMLCSLSAICFFLIGALTPMPAVRTFALTSG 777

Db 373 LSARWSTST--ETPEHRMEETFADAASVITVTSLTDLISFGVGCAATPPSPVQMFCAVAV 430

QY 778 LAVILDFLLQMAFAVALLSLDSKQASRLDVCCC-----KPOELPPPGQEG----- 826

Db 431 AAVVFTYIYQLTFFFAAVMYVTNRREINNR-----HCITPHKLRETLPEKAAKGORSPEK 486

QY 827 --LLLGFFQKAYAPFLHWTIRGVVLLFLFALFGVLSYMSCHISVGLDQELALPKDSYLL 884

Db 487 NSMLAQPFRTTYSDFLLNPLVIRIFILTCFCVYLGASVYCTVKVGLGLENLDPENSYOK 546

QY 885 DYFLFLNRYF-EVGAIVYFVTTLYGYNFSSEAGMNAICSSAGCNFSFTOKI-----QYATE 939

Db 547 RTLMAAEKYFSDYGSSLH---VMMYNLSE---VNVA-----PRKIWNVLEKEVE 589

QY 940 PPEQSYLAIPASSWVDVDFIDW-----LTPSSCCRLYISGPNKOKFCPSTVNSLNCN 992

Db 590 LYEHTEFTAGSDSLRTFLAFVKQAGLLITPE-----NFVYLKN 629

QY 993 CMSITGWSVRPSVEQHFHYLPWFLNDRPNKCPKGGALAYSVNLTSQGVLI-ASRFA 1051

Db 630 -----VFLSQPF-----AKYNRDVVLSDDGEHLEASRI-- 658

QY 1052 YHKPLK---NSQDYTEALRAARELAANITADLRKVPCTDPAFEPFYTITNVFYEQVLT 1107

Query Match	8.1%	Score 562.5	DB 2	Length 1169
Best Local Similarity	21.5%	Pred. No.1.5e-30		
Matches	219	Conservative 186	Mismatch 378	Indels 235
				Gaps 35
Qy	322	GTSLSDKLSFSTHTLLGQPGFGWGTWASWPLTTLVLVSIPVVALAAGLVPELTETDPE	381	
Db	4	GISCVDK-----TLNKSFVHLGICIAKHGPGYFIIPVLVLLTLCMTGYOOLKYQIDPEY	56	
Qy	382	LWSAPNSQARSEKAFHDQHFGPPF--RTNQVILTAPNRSSRYRSDLLGPKNFGSGL---	436	
Db	57	LFSPICEGKTERAIVEQYFKVYTHRFNVCGRITPGR-----FGRVIVIT	102	
Qy	437	-DDLLLELLELQBRLLRHLQWSPEAQNISLQDICYAPLMPDNTSYDCCINSLLQYFQ	495	
Db	103	KXGDNEMIRREVFQELRQLD-----NI-IQN-ATTYDGDVTYKDCNARMENECFE	152	
Qy	496	NNRTLLLLTA-----NOTLMGQTSQVDWKDHPLYCANAPLTKDGTALALSCWAD	545	
Db	153	ND---ILNLDALMDIDRAGQLMLTFFMNFVPTWDAHLF-----	188	
Qy	546	YGAPVFPFLAIGYKGKDYSEAEALIMTFSLNPNYPAGDPLR-----AQAKL	591	
Db	189	---PVP-----FGGTYK-----LTEDNTVISVPALQIVYFTADTKRODAGAE	228	
Qy	592	WEEAFLEMRAPQRMAGMFO---VTFEABRSLEDEINRTAEDELPIFATSYIV--IFLY	646	
Db	229	WEETFLRWVNGAEN--SQQPKHISVSYFASRTLDHELEKNTKTVPYFSSTFLLMGLFSI	286	
Qy	647	ISLALGYSWSRWVNDKATLGLGGVAVVLGAVMAAGPFPSYIGIRSSLVILQVVPFLV	706	
Db	287	ITCMGSD-----AVRSKPFGLMGVNSAIMATLAAFLGAMYCGI-EFIGINLAAPEFLM	338	
Qy	707	LSVGCADNIFIVLVEYQRLPRPGPREVHIGRALGRVAPSMLLCSLEATCFFLGALTPTM	766	
Db	339	IGIGIDTIVMLAGWRTKAK--MPVABRMGLMMSEAAVSTITSVTDFISFLIGISPP	396	
Qy	767	PAVRTFALTSLGAVILDFLLQMSAFALLSLDSKRQBSASRLDVC-CCVQPOELP----	819	
Db	397	RSVRIPCTYSVFAVCFTELWHITFFAACMAISGYRERKNLHISIFGCRVQPMVAIKEBN	456	
Qy	820	-----PGQSGELLGPFQKAYAPPLHLHWITRGVWVLLFLALFGVSL	861	
Db	457	FLYKAIMAGGIDANDPNPIDNKHMLMAFPKDXAAVANNKKCKAIIILAFASLYVGAC	516	
Qy	862	YSMCHISVGLDQELALPKDSVILLDYFLFLNRYPE-----VGAPYFVTTLYGNPSS	912	
Db	517	YGITQIKEGELRRKLSRSDYSVEFFDREDDYREFFYRMOVITAGPL-----NYS	568	
Qy	913	EAGMNAICSSAGCNFFGTOKIQVATEFFBQSYLAIPASSWVDDFIDWLTPSSCCRLYIS	972	
Db	569	PLVQEQV-----ENLTSTLEHTSVTSERY-----TESWLSRFLSFLERN--	609	
Qy	973	GPNKDKFCPSTVNSLNCNKMCSITMGSRVPSVEQFHXYLPWFILNDRPNIKCPKGGLAAY	1032	
Db	610	-----ELLNVTVDDQTFIDAVKEH--WLPPGNP-----P	637	
Qy	1033	STSVNLTSD-GQVLASRFMAYHKLKNSQDYTEALRAARELANITADLRKVPCTDP--A	1089	
Db	638	SLDVRFNEDETOITIASRFLIQAENVITDTHKEKNVR-----DLRQICKDSPLNA	686	
Qy	1090	FEVPPYITNVFEOYLTILPEGLFMLSCLVPTTFAVSCLLGLDRSGLNLLSIWMIL	1149	
Db	687	SIFHPYP---VFPDQFLVRPVSLOAMVIGAIMMIISFVIP-NILCSLWVAFVSISIE	742	
Qy	1150	VDTVGFMALWDISYNAVSLNLSAVGMSVEFVSHITRFSATISKPTWLSRAKEATISM	1209	
Db	743	LGVAGYALMDVNLDSISMINLMICIGFSVDFTAHICVYTMSSKKRSPKARVREALHSLG	802	
Qy	1210	SAVPAGVAMTNLPGLIVGLAKAQIQLPFRMLNLLITLGLLHGLVFLPVLISYVGP	1267	
Db	803	LPITIQSSSTIL-GIVALLAQSYIP-LVFPMKVVFLVFFCAMHGLPFLVLLSLFQF	858	

RESULT 84					
ID	Q4SP00	TETNG PRELIMINARY;	PRT;	816 AA.	
AC	Q4SP00;				
DT	13-SEP-2005	(TREMBLrel. 31, Created)			
DT	13-SEP-2005	(TREMBLrel. 31, Last sequence update)			
DT	13-SEP-2005	(TREMBLrel. 31, Last annotation update)			
DE	Chromosome 15 SCAF14542, whole genome shotgun sequence.				
GN	ORFNAMES=GSTENG00015058001;				
OS	Tetraodon nigroviridis (Green puffer).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Acanthopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;				
OC	Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;				
OC	Tetraodontidae; Tetraodontidae; Tetraodon.				
NCBI_TaxID=99883;					
OX	NCBII_TaxID=99883;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE.				
RA	Jailon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,				
RA	Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,				
RA	Nicaud S., Jaife D., Fisher S., Lutfalla G., Dossat C., Segurens B.,				
RA	Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,				
RA	Anthonard V., Jubin C., Castelli V., Katinka M., Vacherie B.,				
RA	Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,				
RA	Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Guzy J.,				
RA	Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,				
RA	Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,				
RA	Lindblad-Toh K., Birren B., Nusbaum C., Kahn W., Robinson-Rechavi M.,				
RA	Laudet V., Schachter V., Quetier P., Saurin W., Scarpelli C.,				
RA	Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,				
RT	"Genome duplication in the teleost fish Tetraodon nigroviridis reveals				
RT	the early vertebrate proto-karyotype.";				
RL	Nature 431:946-957(2004).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE.				
RG	Genoscope; Whitehead Institute Centre for Genome Research;				
RL	Submitted (PBB-2004) to the EMBL/GenBank/DBJ databases.				
CC	-!- CAUTION: The sequence shown here is derived from an				
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is				
CC	preliminary data.				
DR	EMBL; CAAB01014542; CAP997632.1; -; Genomic DNA.				
SQ	SEQUENCE 816 AA; 90662 MW; 2852CA728580E7D0 CRC64;				
Query Match 7.7%; Score 532; DB 2; Length 816;					
Best Local Similarity 22.0%; Pred. No. 1.4e-28;					
Matches 211; Conservative 160; Mismatches 385; Indels 202; Gaps 30					
Qy	351	WPLTILVLSIPVVVALAAGLVFTS--LTTPDVELWSAPNSQARSKAFHQHGFPPFRTN	408		
Db	13	WRRTAHLSGSDPSRSLGFGFCFLQDIRANDIEEQPTPDVGQAERRIYQB---TFPGN	68		
Qy	409	QVILTAPNRSSRYDSLLLGPNFSGILDLDLLELLEQLQRHLRWQSWPEAQRNISLQ	468		
Db	69	ESMSFRLRLSGSGNYATLIATSE-GDVLTTTVLQDIVELDSEVRMMV--HHDNRSFEYQ	125		
Qy	469	DICVAPLNPDNTSLDYDCINSLIQVFQNNRTLTLTANOTLMGQTSQVDKDHFLVCANA	528		
Db	126	DVCGAVMG-----SCTPNHILDLVGDGR-----AMLTP-----PW-----FHSENG	161		
Qy	529	PLTFPKGTALALS CMADYGAPVPFPFLAIGKYKGXDYSBAEALIMTFSLNNYPAGDPRLAQ	588		
Db	162	SVPLHGG-----LGGVKLWANSSVQSCAIRLFYYLRQQDR-TK	200		
Qy	589	AKLWEERAFLEBMAFORBMACGMFOVTFAERSLEDENIRRTAEADLPATSYIVIFLYIS	648		
Db	201	TDLWLQSFL----RLVSNASSASIRVSYSKSRLQWFQKTGSLCLFSAAYA-----IA	252		
Qy	649	LALGSYSWSKRMVYDSKATIGLGGA VVLGVAMAAMPFSYLGISSILQVWPFLVLS	708		
Db	253	ITPSIVTCNRWDSVRTKKWWALGVGFSTALAVLSGFGALLLLG-RPFVMTAASCCEPLIIG	311		
Qy	709	VGADNIPIFVLVEYOR-----LPRRRGPGEHVHIGRALGRVAPSMLLCSLSAICFFLGA	762		

```
Db 312 VGLDDMFILISCRWRTRVLDSPRR-----LADTYGEAGVSISITTLTNALALLVGY 363
Qy 763 LTPPAVTEALTSLAVILDFLLQMSAFVALLSLDSKROEASRLDVCCVKPQELPPG 822
Db 364 SSPFGSVRSFCLYAGSVCFCLYCVTFGLGACMALNGRREAOQHFTFC-----RKVPDS 419
Qy 823 QGE-----GLLAG-----FFQKAYAPFLLHITRIGVVLFLFALFGV 859
Db 420 AERTNISRICLGGRRGDITEMETEAMTIDIFEKFGYGFLLTHSKVACVLLVAGYLA 479
Qy 860 SLYSMCHISVGLDQELALPKDSYLDYFLFNRYFEVGAPYFVFTTLGYNFSSBAGMNAI 919
Db 480 SIYGLILKEGLEIKNLVLDSDSIIPYLEDQKHF-----GEYGFNV 522
Qy 920 CSSAGCNFSTFKIQYATEPEQSYLAIPASSWVDDFIDMLTPSSCCRLYISGNKDKF 979
Db 523 -----VAKQPLLYWQSEQ-----RLH-----S 542
Qy 980 CPSTVNSLNLKNCMSITMGSVR-----PSVEQPHKYLPMFLNDRPNIKCPKGGLA 1030
Db 543 CVSRFEGTLFVSGTILSWFLSQRYSNTNRLDVTSAERTHLTHLEANP-----592
Qy 1031 AYSTSVNLTSGQVLASRFMAHKLKNSQDYTEALR-AARELANITADLRKVPGTDP 1089
Db 593 VFKQDISFSAGQIQASRFFITQLQIPKEEVMVTLRQTABECPQLLLV-----FHPA 645
Qy 1090 FEVP-PYITITNVFEQYLTILPEGLFMLSCLIVPTFAVSCLLGLDLSGLNLLSIWMI 1148
Db 646 FYLDQYTVVTAKTQVAVLAVALVLSLALIPSLCSAW-----AFVCSV 694
Qy 1149 LVDTVGFMALWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLBRAKEATISM 1208
Db 695 MGVAGFVALMGVNLDSVMINLVCTGFSVDFSAHVSFAVSSSKTDIDGKAVELARL 754
Qy 1209 GSAVFAGVAMTLPGLVILGLAKAQLIQIFFRMLNLTLLGLHGLVFLPVILSYVG 1266
Db 755 GYVPLQG-ALSTILGWLLSGSVIFRTFFKIFLVIT-FGLHSLVFIPIVFLTLG 810

RESULT 85
Q9XYP5 JUNCO PRELIMINARY; PRT; 1318 AA.
AC Q9XYP5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative hedgehog receptor.
GN Name=ptc;
OS Junonia coenia (Peacock butterfly) (Precis coenia).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Papilionoidea; Nymphalidae; Nymphalinae; Junonia.
OX NCBI_TaxID=39708;
RN [1]
RL NUCLEOTIDE SEQUENCE.
RP MEDLINE=99115913; PubMed=9915699; DOI=10.1126/science.283.5401.532;
RA Keys D.N., Lewis D.L., Selegue J.E., Pearson B.J., Goodrich L.V.,
RA Johnson R.L., Gates J., Scott M.P., Carroll S.B.;
RA "Recruitment of a hedgehog regulatory circuit in butterfly eyespot
RT evolution.";
RL Science 283:532-534 (1999).
DR EMBL; AF117898; AAD31595.1; -; mRNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR004392; Patched.
DR InterPro; IPR004766; Patchedtm_recept.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR TIGRFAMs; TIGR00918; 2A060602; 1.
DR PROSITE; PS50156; SSD; 1.
KW Receptor.
```

```
SQ SEQUENCE 1318 AA; 146474 MW; BA774C7133A22221 CRC64;
Query Match 7.6%; Score 522.5; DB 2; Length 1318;
Best Local Similarity 22.3%; Pred. No. 1.2e-27;
Matches 259; Conservative 173; Mismatches 421; Indels 307; Gaps 47;
Qy 336 ILGQFFQGWGTWASMTLILVLSVIPVVALAAGLVFTELTDPVELMASPNSQARSKA 395
Db 72 ILGCLQGDAGKVLV--VAIIVLSTFCV-----GLKSAQIHTRVDQVWVGGRLEAE 124
Qy 396 FHDQHPFP-PRTNQVILTPNRRSYRDSLLLPKPNFSGILDLDDLLLELLELQERL 454
Db 125 YTAQALGEADSSHTQLVI---QTAQKPDVSLHP---GAL-----LHLL 162
Qy 455 QVNSPEAQRNI-----SLQDICVAPLNDP-----NTSLYD-----CCINSL 497
Db 163 KVVHATRTVTHMYDIEWRUKDLCISPIDFEGYHHIESIIDNVIPCAITPDLDCF 222
Qy 498 RTLL-----LLTANQTLMGQTSQ---515
Db 223 SKLIGPDYPIVPHLKHKLQWTHLNLPLEVVEVKLKFQPPPLSTIEAYMKRAGITS 282
Qy 516 ---VDWKHFLYC-ANAPLTFKQGTALALSCHADYGAIVPFF-----LAIGYKGD 563
Db 283 KPCLDPTDP--HCPATAP-NKSGHI PDVAEALSHGCGYFAAAYMHMPEQILVGGAT 339
Qy 564 YS---EAEALIMTFSL-----NNYPAGDPRLAQAKLWEBAFLBEMRAFORRMAG 614
Db 340 TSAUSARALQTVVQLMEREMYWADYKHQIIGNQEKAAAVLDQWQKFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTTAEDLPFATSIVIVFLYISIALGSYSWSRV 661
Db 400 TTSGSVSSAYSFYFSTSTLNDILGKFEVSLKNIILGYMFLIYVAVTL---IQW-RDP 455
Qy 662 VDSKATILGLGVAVVLGAVMAAMGFFSYGLTRSSILVILQVFPFLVLSVGADNIF 721
Db 456 IRSQGVGIAGVLLSITVAAGLGCALLGFPFNASTQIVPFPALGLGVQDMFLTH 515
Qy 722 QRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEACFPFLGALTMPMPAVRTALT 781
Db 516 ---VEQAGDVPREERTGLVKKSGLSVLLASLCNVMAFLAALLPIPAFRVCLQA 573
Qy 782 LDFLQMSAFVALLSLDSKROEASRLDVCCVKPQELPP-----820
Db 574 FNLGSILLVFPAMISLDRRESAARADLLCCLMPESPLPKKKI PERAKTRKNDK 633
Qy 821 -----PGQGEGL-----LLGFFQKAYAPFLHITRIGVVLFLFALFGV 865
Db 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILT 693
Qy 866 HISVGLDQELALPKDSYLDYFLFNLR---YFEVGAPVYFVTTLGYNFSSBAGMNA 922
Db 694 KVQGLDLTDIVPENT---DEHFLSRQEKYFGP-----YN-----MYAVTQ- 732
Qy 923 AGCNNSF--TQKIQYATEPEEQSYLAIP-----ASSWVDDFIDWL----- 961
Db 733 ---GNFEYPTNQKLLY--EYHDQ-FVRIPTNIKNDNGGLTKFWLSLFRDWLLDL 786
Qy 962 -TPSSCCRLYISGNKDKFCPSTVNSLNLKNCMSITMGSVRPSVEQ-----1007
Db 787 EVASGCI-----TQEWYCKNASDE-GILAYKLMVQTGHVDNFDKSLITAGH 838
Qy 1008 -----FHKYL-PWFLNDRPNIKCPKGL-----AAYSTSVNLTSGQVLAS 1050
Db 839 GIINPKAFYNLSAWATNDALAYGASQNLKPKQQRWHTSPEDVHLEIKSSFLY 898
Qy 1051 AYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAFEVFPTITNVFEQYLT 1110
Db 899 FYLSGLSDTDSIKTLIRSVRDLCKYEA--KGLPN-----FPGSIFPLFWEQYL 948
Qy 1111 EGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFMAWDSYNVSLIN 1170
Db 949 TSLILALACALGAVPIAVVLLNNAWAAVLTALATLVQLLGVNALLGVKLSAMP 1008
```

```

QY 1171 LVSAGMSVEFVSHITRISFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGLVLGLA 1230
D 1009 LVLAIGRGVHFVHLCGLFVTSIGC---KRRASIALES-VLAPV---VHGALAAALA 1059
QY 1231 KAQL-----IQIPFRLNLTLLGLLHGLVFLPVLSYGP--DYNPALALEQKRAE 1281
D 1060 ASMLAASEFGFVARLFLRLLLALVFLGLIDGLFFPVLISLGPAAEVRPI----- 1110
QY 1282 EAVAAMVASCPNHPFSRVST 1301
D 1111 -----EHPERLST 1118

RESULT 86
ID Q61232_CABER PRELIMINARY; PRT; 1402 AA.
AC Q61232_
DT 25-OCT-2004 (TRENBLrel. 28, Created)
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)
DE Hypothetical protein CBG03294.
GN Name=CBG03294;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI TaxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000012; CAE59824.1; -; Genomic_DNA.
DR GO; GO:0036020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR00731; SSD 5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS50156; SSD; 1.
KW Hypothetical protein.
KW SEQUENCE 1402 AA; 155837 MW; 75584DEA4679D650 CRC64;

```

```

Query Match
Best Local Similarity 7.6%; Score 522; DB 2; Length 1402;
Matches 278; Conservative 178; Mismatches 435; Indels 406; Gaps 55;

QY 321 KGTSLSKLSFSTHTLLGQFFQGWGTVV--ASWPLTILVLSVPVVALAAGLVFTLTDT 378
D 100 RKGATGRVALYSRSLIQKLLFALGNVHNAWSI-ILVWSMI-PAFCYGLQYVHIETD 157
QY 379 PVELWSAPNSQARSKAP-----HDQHFQFPF-RTN-----QVILTAPNRSYR 421
D 158 IVKLVAQGRLEDEELNLPNMKELMRNRESVDSGPEIPRENGLGGYQVLIQTPEFEG-- 215
QY 422 YDSLGLGPKNFGILDLDLLELLELQERLHLQWSPEAQNI-----SLQDICYP 474
D 216 QDALAQGP-----LLKHVEIMKHIAFSNVSVHGVDSLSLSDICPKP 255
QY 475 ----LNPDN--TSLYD-----CC-INSLIQYFQNNRTL-----LLILTA 505
D 256 APPSVSPDSPAASLSDVIDKIVPCWITPIDCFWEGSKALGPHPSLPKSLGLVGMLLKA 315
QY 506 NOTLMGQTSQVDWKD-----HFLYCANA-----PL----- 530
D 316 ----LPDHDMIRWSDFPIKVIDEITHNSFNLGSHYTFPERAGVGHYNNRPPCIDPLDPEC 371
QY 531 -----TFKQCTALASCWADYGAPVPPF----- 554
D 372 PPMKRYFDVCHVDKIREVASKYQGLEQEKQKDTGDFVFSFGKRETDQQPQMIHP 431
QY 555 -----AIGGY----- 559

```

```

Db 432 AQPADSPVVEGAVPADIPVTADPVTVTTLTSSREVENAKKIKOREMREHCKSYRKSTFN 491
QY 560 -----KGDYSEAEALJMT-----FSLN--NYP-----AGDPR----- 585
D 492 WLKANQDKWTEVPENVPKNDVYAE-----MTGCSGFASNVLNWPMEDMILGNPHRTKR 547
QY 586 ---LAQA-----KLMEAFLEE-MRAFORRM 607
D 548 GKLTSDALQSFLVASPADVFLRPFKQKQGRNPPKSLDQEAWNETAABEQVLQAWQNF 607
QY 608 AGMP---QVTTAE-----RSLEDEINRTAEDLPIPAT-SYIVIFLYISLALGSYSSWR 659
D 608 TKSLYTHEANFDADGNERRTLHPLASTSIADMLBEFCQFNYYTIIIFAGVALML-AVAITIQ 666
QY 660 VMVDS--KAT-----LGIGGVAVVLGAVMAAGPPSYLGIRSSLVILQVVPFLVLSVGAD 712
D 667 ARFDSCLPATSSMGLALAGVLVVTFSVAGVLATWFGIEFNAATQIVPELTIGIGVD 726
QY 713 NIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFILGALTMPMPAVRTF 772
D 727 NMFMLLHNYRDVVKLAGGHAEMAI--LMRETGMSILCTSINNLSFLTGTLLPIPALRSF 784
QY 773 ALTSGLAIVLFLLOMSAFVALLSDSKQASRLDVCCV----- 813
D 785 CAQSSILLTFNFIAITTYPAIISIDLRKKAQRDLCLCYGDTREESYSMISKPKIPS 844
QY 814 ----KPOELPPPGQEGE-----LLGFFQKAYAPFLHWHITRGVVLPLF 853
D 845 KRIIGAPSEASIMQOFGDITQAOWASSDDPAPSLHAFIRYIYIPFISRPACKVGIIGC 904
QY 854 LALFVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFVEVGAPVVFVTLTGVNFS-- 911
D 905 LVLLAASYGMOQSTLGLSLGDLVPEHTAPQFLRDKYFSP-YPM-FAVIKGNIDYA 962
QY 912 -SEAGMNAICSSAGCNFSFTQKIYATFEFPEQSYLAIPASSWDDFIDWLPSSCCRLY 970
D 963 HQQRQIDNYROSIGSSKEFVKNK---NEEPSEKY-----WLGIMRDLIT--STQGRF 1009
QY 971 ISGPNKDKFCPST-----VNSLNC-----LKNCMSITWGSVR-----PSV 1005
D 1010 DEEVAKGSFDLVGTIVIAANVSEDAIAHAIHMLCSHGELYECAG-RVGKIRLVDAAGIINS 1068
QY 1006 EQFHXYL--PWFINDRPNIKCPKGLAAYSTSNLTSDQVLASRSMAYHKPLKNSQ--DY 1062
D 1069 DGFNYLTAMFNVDHMYTVVSQASFPPTPKWELVKNK---ENFIPRAKPLVYSQIPY 1125
QY 1063 TEALRAARELANITADLRKYPG--TDPAFVFPVPTITNVFYEQYLTILPEGLFMLSCL 1120
D 1126 LTGLTDTGVIVEAI-KDIRSVCDRPTDQGLPNFPQGIAPTEQYLFELTGNLMQAISIT 1184
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIVMTLVTVGFMAWDISYNAVSLINLVSAGMSVE 1180
D 1185 ISVFCVISVLL--FNPWAAALMVVCTILGIMTCLAGFMGLVGLKINPVSATVITAVGIGE 1243
QY 1181 FVSHITRSFAISTKPTWLERAKEATISMGSAVAGV---AMTNLPGLVLGLAKAQLQI 1237
D 1244 FTVHVSFL-----TALCTRAQRTSSAVDRVFPVVIHGSFSTLLGILMLGSEFVVK 1298
QY 1238 FFFRLNLTLLGLLHGLVFLPVLSYGP--DVNPA 1272
D 1299 YFFIVMTALIGTINGIINGILLPVLLSWFGPRREISPA 1335

RESULT 87
QY Q4RVD6_TETNG PRELIMINARY; PRT; 1197 AA.
ID Q4RVD6_
AC Q4RVD6_
DT 13-SEP-2005 (TRENBLrel. 31, Created)
DT 13-SEP-2005 (TRENBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TRENBLrel. 31, Last annotation update)
DE Chromosome 15 SCAF14992, whole genome shotgun sequence.
DE (Fragment).

```


RX MEDLINE=90058658; PubMed=2582494; DOI=10.1016/0092-8674(89)90021-4;
RT Hooper J.E., Scott M.P.;
RA "The Drosophila patched gene encodes a putative membrane protein
RT required for segmental patterning.";
RL Cell 59:751-765(1989).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90015164; PubMed=2797178; DOI=10.1038/341508a0;
RA Nakano Y., Guerrero I., Hidalgo A., Taylor A., Whittle J.R.S.,
RA Ingham P.W.;
RT "A protein with several possible membrane-spanning domains encoded by
RT the Drosophila segment polarity gene patched.";
RL Nature 341:508-513(1989).
RN [3]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.H.C., Blasej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mentrulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [4]
RP GENOME REANNOTATION.
RX MEDLINE=22426669; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [5]
CC -!- FUNCTION: Segmentation polarity protein. Acts as a receptor for
CC the hedgehog protein (HH). Associates with the smoothened protein
CC (SMO) to transduce the hedgehog signal leading to the activation
CC of wingless, decapentaplegic and patched itself. Participates in
CC cell interactions that establish pattern within the segment and

CC the imaginal disks during development. In the absence of HH,
CC represses the constitutive signaling activity of smo through fused
CC (FU).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the patched family.
CC -!- SIMILARITY: Contains 1 SSD (sterol-sensing) domain.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC
CC EMBL; M28999; AAA28696.1; -; Genomic DNA.
CC EMBL; M28418; AAA28696.1; JOINED; Genomic DNA.
CC EMBL; X17558; CAA35591.1; -; mRNA.
CC EMBL; AE003836; AAF59062.1; -; Genomic DNA.
CC PIR; S06119; S06119.
CC FlyBase; FBgn0003892; ptc.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0008158; F:hedgehog receptor activity; IMP.
CC GO; GO:0004888; F:transmembrane receptor activity; NAS.
CC GO; GO:0004899; P:anterior/posterior lineage restriction; TAS.
CC GO; GO:0001746; P:Bolwig's organ morphogenesis; IMP.
CC GO; GO:0009880; P:embryonic pattern specification; TAS.
CC GO; GO:0007455; P:eye-antennal disc metamorphosis; IMP.
CC GO; GO:0042078; P:germ-line stem cell division; TAS.
CC GO; GO:0004877; P:negative regulation of smoothened activity; TAS.
CC GO; GO:0030707; P:ovarian follicle cell development (sensu in. .); TAS.
CC GO; GO:0007422; P:peripheral nervous system development; TAS.
CC GO; GO:0007346; P:regulation of mitotic cell cycle; IMP.
CC GO; GO:0042306; P:regulation of protein-nucleus import; IMP.
CC GO; GO:0007224; P:smoothened signaling pathway; IMP.
CC GO; GO:0048103; P:smoothed stem cell division; TAS.
CC GO; GO:0048100; P:wing disc anterior/posterior pattern formation; TAS.
CC InterPro; IPR0003392; Patched.
CC InterPro; IPR000731; SSD 5TM.
CC Pfam; PF02460; Patched; I.
CC TIGRPFAMs; TIGR00918; 2A060602; 1.
CC PROSITE; PS50156; SSD; 1.
CC Developmental protein; Glycoprotein; Receptor;
CC Segmentation polarity protein; Transmembrane.
CC TOPO DOM 1 76 Cytoplasmic (Potential).
CC TRANSMEM 77 92 Potential.
CC TOPO DOM 93 427 Extracellular (Potential).
CC TRANSMEM 428 448 Potential.
CC TOPO DOM 449 465 Cytoplasmic (Potential).
CC TRANSMEM 466 486 Potential.
CC TOPO DOM 487 492 Extracellular (Potential).
CC TRANSMEM 493 511 Potential.
CC TOPO DOM 512 532 Cytoplasmic (Potential).
CC TRANSMEM 533 553 Potential.
CC TOPO DOM 554 562 Extracellular (Potential).
CC TRANSMEM 563 583 Potential.
CC TOPO DOM 584 677 Cytoplasmic (Potential).
CC TRANSMEM 678 699 Potential.
CC TOPO DOM 700 931 Extracellular (Potential).
CC TRANSMEM 932 952 Potential.
CC TOPO DOM 953 955 Cytoplasmic (Potential).
CC TRANSMEM 956 976 Potential.
CC TOPO DOM 977 1007 Extracellular (Potential).
CC TRANSMEM 1008 1028 Potential.
CC TOPO DOM 1029 1056 Cytoplasmic (Potential).
CC TRANSMEM 1057 1077 Potential.
CC TOPO DOM 1078 1082 Extracellular (Potential).
CC TRANSMEM 1083 1103 Potential.
CC TOPO DOM 1104 1286 Cytoplasmic (Potential).
CC DOMAIN 428 SSD.
CC COMEBIAS 1227 1332 Poly-Pro. (GlcNAc. .) (Potential).
CC CARBOHYD 142 142 N-linked (GlcNAc. .) (Potential).
CC CARBOHYD 298 298 N-linked (GlcNAc. .) (Potential).

FT CARBOHYD 335 335 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 388 388 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 807 807 N-linked (GlcNAc...) (Potential).
FT CONFLICT 111 111 R -> G (in Ref. 2).
FT CONFLICT 274 274 G -> A (in Ref. 2).
FT CONFLICT 332 332 A -> R (in Ref. 1).
FT CONFLICT 636 636 P -> A (in Ref. 2).
FT CONFLICT 862 864 DVF -> ASSPTELLRANCIRNR (in Ref. 2).
FT CONFLICT 866 866 Y -> N (in Ref. 2).
SQ SEQUENCE 1286 AA; 5P22A956F8BE0EC6 CRC64;

Query Match 6.9%; Score 474; DB 1; Length 1286;
Best Local Similarity 20.4%; Pred. No. 3.2e-24;
Matches 245; Conservative 177; Mismatches 457; Indels 322; Gaps 42;

QY 321 KGTSLSDKLSFSTHTLGGFCGWTGVAWSWPLTILVLSVIPVVALAAGLVFTELTTPDV 380
DB 43 KKGARGRTAIYLRVPSQSHLETGLSSVQKHAGLVFVAILVLSFVCGLKSAQIHSKVH 102
QY 381 ELWAPNSQARSEKAFHQHFG-PFRTNQVILTA---PNRSSYRYSLLGPKNFSGL 436
DB 103 QLWIOGGRLEALAYTKTIGEDSATHQLLIQTTHDPNASVLPQALL-----A 153
QY 437 DLIDLELLELLELRLHGVMSPEAQRNLSLODICVAPLNPNTSLY-----DCC 486
DB 154 HLEVLVKATAVKVHLTYDE-W-----GLRDMCMPTSPFEGIYIEQILRLHLP 204
QY 487 INSLLOYFQN-----NRTLLLLLTAN-----QTLWGQTSQVDWKDH 521
DB 205 IITPDLDFWEGSGLGPESAVVPGNLQRLMTLTPASVMYQMKQSEKISDFETV 264
QY 522 FLYCANAPLTFKDGTA-LALSQW-----ADYCAPVFPFLAIG--GY 559
DB 265 EQMKRAAI-----GSGYMEKPCLPNPNCPDPAKNKSTQPPDGA---ILSGCYG 316
QY 560 KGQYSEBALIMTSLNNYPAGDRLAQA-----589
DB 317 AAKHMHPEELVGGAKRN-RSGHLKQAQALQSVVQLMTEKEMYDQWQNYKVHHLGW 375
QY 590 -----KLWEEAF--LEEMRAFORMAGMQVFTFAERSLEDEINRTTAEDLPI 639
DB 376 EKAAEVLANQNRNFREVEQLRKQSRITATNDIYVFSAAUDDILAKFSPHSALSIV 435
QY 640 YIVFLYISLALGSSYSSSRVMDSKATILGLGVAVVLGAVMAAMGFYSILGIRSLVIL 699
DB 436 VAVTVLY---ACITLLRW-RDPVGGSSVGVAGVLLMCFSTAGLSALLGIVFNA 491
QY 700 QVVPFLVLSVGNADNIFIVLEYORLPRRPGPREVHIGRALGRVAPSMMLCSLSBAIC 759
DB 492 QVVPFLALGLGVDHIFMLTAAYAESNR--EQTKL---ILKKVGPSILFSACSTAGS 545
QY 760 LGALTPMPAVRTPALTSGTGLAVILDFLQMSAFVALLSLSDSKQEAASRLDV-CCCV 813
DB 546 AAATFVPVPAKVFCLQAAIVMCSNLAAALLVPAMISLDLRRRTAGRADIFCCCF 605
QY 814 KQELPPL-----PGQEGEL-----LLGF 831
DB 606 QPKVAPPVLPNNNGRGARHPKSCNNRVPLPAQNPLLEQRADIPGSHSLASLSATP 665
QY 832 FOKAYAPFLHWTIRGVVLLFLALFGVSLVSMCHI SVGLDQELALPKDSYLLDFLFL 891
DB 666 AFQHTVFPFLMRSWKFLVWVGFLAALISLSYASTRLQDGLDIIILVPRKDSNEH 725
QY 892 RYFEGVAPVFTTLYGYNFSSAGNAICSSAGCNFSTQKIQVATFPEGSYLAIP-- 949
DB 726 RLFL--GFYSMYAVTQ-----NFEYPTQQQLRDY-HDSFVRVPHV 763
QY 950 -----ASSVVDIDMLTPSSCCRLYISGPNKDKFCPTVNSLNCNLC- 993
DB 764 IKNDNGGLPDFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWFFNASSD 811
QY 994 -----MSITMGSVRPSVEQ-----FHKYL--PWFILNDRPNKCPKG 1027

DB 812 AILAYKLIVQTHGVNPNVDKELVLNRLVNSDGIINQRAFYNYLSAMATND----- 862
QY 1028 GLAAYSTSVNLTSDQVNLASRFMAYHKPKNSQD-----YTEALRAARE 1071
DB 863 -VFAVG-----ASQGLYPEPRQYFHP--NEYDLKPKSLPVVAQMPFYLHGLTDSQ 914
QY 1072 LAANTADLRKVPDTPAFEV--FFYTITNVFYEQYVLTILPEGLPMLSLCLVPTFAV 1129
DB 915 1-KTLIGHIRDLVSVKYGFGFLPNYPSGIPPIFWEQYMT-LRSSLAMILACVLLAALV 972
QY 1130 LGLDLRSGLLNLISIVMLVDTVCGFMALWDISYNAVSLINLVSAVMSVPEVSHITRSF 1189
DB 973 LLLSVAAVILSVLASLAQIFGAMTLLGKLSAIPAVILISVGMMLCF--NVILSL 1030
QY 1190 AISTPTWLERAKEATISGSAVFAVAMTNLPGILVLGLAKAQITQIFFEFLNLITLL 1249
DB 1031 GPMTSVGNRRQVQLSMQSLGPLVHGMLTSGVAVFMLSTSPFVIRHFVCLLVVLCV 1090
QY 1250 GLLHGLVLPVLTVSVGPDVNPALALOKRAEAAVAVMVASCNPHSPSRVSTADNIYV 1309
DB 1091 GACNSLLVFPILLSVVGPEAB-LVPLE-----HPDRISTPSPLPVRS 1131
QY 1310 S 1310
DB 1132 S 1132

RESULT 90
Q09938 CABEL PRELIMINARY; PRT; 1015 AA.
AC Q09938;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein C54A12.1.
GN Name=ptr-6; ORFNames=C54A12.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoida;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for investigating biology."
RL Science 282:2012-2018 (1998).
DR EMBL; U28733; AAA68302.1; -; Genomic_DNA.
DR PIR; T15830; T15830.
DR Ensembl; C54A12.1; Caenorhabditis elegans.
DR WormBase; WBGene0004221; ptr-6.
DR WormPep; C54A12.1; CE01865.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1015 AA; 116343 MW; 6D69AF443028F6C5 CRC64;

Query Match 6.9%; Score 473.5; DB 2; Length 1015;
Best Local Similarity 19.2%; Pred. No. 2.6e-24;
Matches 226; Conservative 186; Mismatches 417; Indels 347; Gaps 38;

QY 289 IILCSVPAV---VTILLVGRVAPARDKSKWDPKKGTSLSKLSFSTHTLGGFCG 345
DB 11 LLKTSIFENGKNVLLALIAHRF-PNRCRIPFLDKK---LEKVFANYTKHLLVDY 61
QY 346 TWVASWPTILVLSVIPVVALAAGLVF-TELT-----DPV-----ELWSA 385
DB 62 -----YVFLIAPVILTACSGFYWKELTLLDARKLYTPVSAPSWKEQVFS 114

```
QY 386 PMSQARSEKAFHQHGFPPRTNQVILTPNRSYRYDSLLGPKNFGSGLDLDLLELL 445
Db 115 KSYEFLPERT-----FQNRXYLVVHGRPYENG-----YFNILEGSLYNDIA 158
QY 446 ELQERL-----RHLQWSPQAQRNLSLQDICYAPLNPNONTSLYD 484
Db 159 KIEBELVTNVPFREALEDNPAWKRKHLETEIEENRITI-----TFKD 205
QY 485 CCINSLGYFQNNRTLLLTANQTMGQTSQVDWKDHFYCANAPLTFKDGTAALSCMA 544
Db 206 VCMWYGCYREKNVIELKRHELN-----KRGISVTFPOVN 243
QY 545 DYGAPEVFPPLAIG--VKGRDYSABEALIMTFSLNNYPAGDPRLAQAKLWE---EALFE 598
Db 244 QEGPIYIAFVIGGVDTPNDTIKARAWRLWYFLKFDDEQEQL--AKFWEDTAEREVR 301
QY 599 EMRAQFRMAGMFQVTTAERSLEDEINRTTAEDLPIFATSYVIFLYISIALGSYS--- 655
Db 302 ETYADHPTI---QCHIKHSRIVDOGLTENANRLKPYENVITAVLVLFAP-----YSVKW 353
QY 656 -----SWSRWVD---SKATLIGGGVAVVLGAVNAWGFYSYLGIRSLVLQVVPVLV 706
Db 354 YFRMDHWP-LHIDWLRKPMALGGVLSVLAILSGIGLLWFGMFFAEITL-IAPFLV 411
QY 707 LSVGADNPIFVLEY---QRLPREPGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGA 762
Db 412 LSGVDDMFIAVAWHNTEMYPGSPKMKRMTEAMSESAVAIFITSFIDVLSFGAGT 471
QY 763 LTPMPAVRTFALTSGVLVILFLLQMSAFVALLSLDSKRQASRLDVCCV----- 813
Db 472 LTDITAVOGFCAMTAACMFFFLYQITFFAALMVISAQAQSGRNSCMPCITAGDIYIE 531
QY 814 -----DQELUPP 820
Db 532 DGSQPNLKKKSKSRDKAKAEKKEKBAKNDKNMEIEBNAENAEKSSYDSDPPSQIHI 591
QY 821 PQCGELLLGFQKAYAPLHWHITRGVVLHLLFLALPGVLSYSMCHISVGLDQELALPKD 880
Db 592 PVKSRGAMGHFRDYVVPWLLNWKTKLPMFVTFIYLGISVVICMVEQGLDYDKLLHS 651
QY 881 SYLLDYFLFNRYFVGAPVYVVTTLGYNFSEAGMNAICSSAGCNFSTQKIQAYATEF 940
Db 652 DELVE---ALKXEIEL-----FHGGQDIEA--- 674
QY 941 PQSYLAIPASSWVDVDFDMLTPSSCCRLYISGPNKDFCPSVNSLCLNKMCSITWGS 1000
Db 675 -----IQNCPNITIAE 685
QY 1001 VRPSVEQFHKYLPLWFLNDRPNIKCPKGG-----LAAYSTSVNLTSD-----GQ 1043
Db 686 SDRDIE-----LVAQFENISYSLGGKGTGFWREYKYSNLTGSYLNDRNRESNIVGV 738
QY 1044 VLASRFMAYHK-----PLKNSQDY-TEALRAAR-----ELAANITADLRKVP 1084
Db 739 YWMSQLFAYKLSQDFWANESDYDTLELKSRYFRIGVHRLSTDTDLV-ITEELRGA 797
QY 1085 GTDPAFEVPPYITNVFYEYQYITLPEGLFMLSCLCLVPTFAVSCLLGLDRLSGLLNLS 1144
Db 798 DRHPDLNIVTYQOSRAIADQLNVLSSTITNDLAFMFCVALLFIPNPI-CALFITFA 856
QY 1145 IVMILVDVTGFMALWDISVNASVLINLSVAGMSVEFVSHSTRSPAIS-TRKPTWLERAKE 1203
Db 857 MYTIDIGVIFUSLSVKLDPISMITIIIMSIGFISIEFSAHIHFGVFNDSNLSAFDRCDV 916
QY 1204 ATISMSAVFAGVAMTNLPGIILVGLAKAQLTQIEFFRLNLIITLLGLLHGVLFPVLIS 1263
Db 917 AMEKLAWPVHG-SLSTILGVFLAFIDSVMVLVFFKTLISLV-LIIGAWHALMLLPILLS 974
QY 1264 YVGP-----DVNPALALEQKRAEBAVAVWVASC 1293
Db 975 MCPIVIERLSDASKASDRRRKLKENQNSVYAINLP 1010
```

```
RESULT 91
Q61PCI CAEBR
ID Q61PCI_CAEBR PRELIMINARY; PRT; 915 AA.
AC Q61PCI;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein CBG07607.
GN Name=CBG07607;
OS Caenorhabditis briggsae.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxID=6238;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RG The C.briggsae Sequencing Consortium;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DDAJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAC01000032; CAB63242.1; -; Genomic DNA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; P:hedgehog receptor activity; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD_5TM.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Hypothetical protein.
SQ
SEQUENCE 915 AA; 103951 MW; DD7AB4E5B0125E50 CRC64;
Query Match 6.8%; Score 470.5; DB 2; Length 915;
Best Local Similarity 20.2%; Pred. No. 3.6e-24;
Matches 208; Conservative 195; Mismatches 391; Indels 235; Gaps 32;
QY 336 LLGQFPQGWGWASWPLTILVLSVIVPVALAAGLVFTE--LTDPVVLWSAPNSQARSE 393
Db 9 LVRNAFEGYGVVHRWFCFISPLFTMACSIGFLRWKDLRVDDPAYVFPFSDARWRE 68
QY 394 KAFDQHF-----GPFRTNQVILTAPNRSYRYDSLLGPKNFGSGLDLDLLELL 444
Db 69 INVFNENWFLDNKFLPGKSPETK-----RVNIIIRAKDGGSIMRDNVLOBI 116
QY 445 LELOB-RLRHLQWSPQAQRNLSLQDICYAPLNPNONTSLYD--CCINSLGYFQNNRTLL 501
Db 117 QILANNIMNNTSVPTDLDLKFNLTYQDLCLS-----YDWVCGANBHIQMLIRNDV- 166
QY 502 LLTANQTMGQTSQVDWKDHFYCANAPLTFKDGTAALSCMADYG-APVFPFLAIGGYK 560
Db 167 ----NQVLDLHFPFGTKYKTPVYLS---IFGDVNFENGTLSDAKLTQLFYFLQDQKM 219
QY 561 GKDYSEALIMTFSLNNYPAGDPRLAQAKLWEAEFLBEMRAFORRMAGMFOVTTAERS 620
Db 220 VEEYSKFSYAIEFLNQVSSD-----VISLSFAHYQS 253
QY 621 LEDEINRTTADDLPIFATSYVIFLYISIALGSYSSWSRVWD---SKATLGLGVAVVL 677
Db 254 LEDGLDENAKAFVNFVFFVFLAMAYALISSFLSKSSNAKKIDMTISSKFWLATAGMTFTV 313
QY 678 GAVMAWGFPPSVGLIRSSLVLQVVPFLVLSVAGDNIFVLEVORLPRRCPGEPRVHIG 737
Db 314 LSVSAFGLFLLGVHYN-VINTIIPFLIIAIGDDMLMNAWQDTSKLSVPE--RMS 370
QY 738 RALGRVAPSMLLCSLSEAICFFLGPALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSL 797
Db 371 KTLSHAGVAVITNTVDIMSFAIGCITDLPGLQPFQCIACVSVAFSYLYQLTFPSGMAI 430
QY 798 -----DSKRQE----- 803
Db 431 MGEVEREKRHCLFFYRTLRLVDISKONEADSKLQOIKNRSPASPAPFNLYSSNSFSFSD 490
```

QY 804 ----ASRLDVCCVKQELPPPG-----QSEGLLGFQKAYAPFLHWT 845
 Db 491 TDSFSKKTIPAEFAWKQEPSNSSISEKSKSHGPHBEKDRIVYFLGQVYGFILNSV 550
 QY 846 RGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYEFGAPVFT- 904
 Db 551 RLPSAVIFALYVIALYGCNREGNPGNLTNDHYIAKYSDIXHFRIGPOLHVAVL 610
 QY 905 -----TLGYNFSEAGMNAICSGAGCNFSFTQKIQYATEFPEQSYLAIPASSWDDFD 959
 Db 611 NPKLTLGEN--REKLLDVVKS---PENTPYTLG-REGTVF-----FLLEYVN 652
 QY 960 WLPSSCCCLYISGPNKDFCSTVNSLNCNCSITMGSVRPSVEQFHKYLPHFLNDR 1019
 Db 653 YLTD-----LNVEIADTEKLMNTKLS-----WLKYTGSGTQ-----WASNIR 690
 QY 1020 PNKCPKGLAAYSTSVNLTSQOVLASRFMAVHKPLKNSQDVTEALRAARELANITAD 1079
 Db 691 MN-----GTEFOSFRPQALKNFVFPNDHKHAKLLRIAD----- 726
 QY 1080 LRKVPGTDPAFVFPVYITNVFYEQYLTLPL---EGLFMLSCLVPTFAVSCLLGLDLR 1136
 Db 727 -RQYNVNVVYHEAFP-----FADQYLIIIPATIQNVVISLLCM---AVVSFLVPSLP 775
 QY 1137 SGLNLLSIVMLVDVTPGMAWDISYNAVSLINLSVAGMSVEFVSHITRSPASTKET 1196
 Db 776 SGFVFSVSVSINIGVGYMTLWGNLDVSMISIMSGFADLSAHIYAFVTSHGDT 835
 QY 1197 WLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLQIIFERLNLITLGLLHGLV 1256
 Db 836 -KQVIGALETLQWPIFOGASST-IGISILYTVDAYII-LVFFKTIWMLTGLAIGHLP 892
 QY 1257 PLFVILSYV 1265
 Db 893 FIPVFLSVV 901

RESULT 92

Q621T1 CAEBR PRELIMINARY; PRT; 974 AA.
 AC Q621T1 CAEBR PRELIMINARY; PRT; 974 AA.
 DT 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Hypothetical protein CBG02358.
 GN Name=CBG02358;
 OS Caenorhabditis briggsae.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxId=6238;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RG The C. briggsae Sequencing Consortium;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 CC -! CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

DR EMBL; CAAC0100012; CAES9072.1; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
 DR GO; GO:0007275; P:development; IEA.
 DR InterPro; IPR003392; Patched.
 DR InterPro; IPR000731; SSD 5TM.
 DR Pfam; PF02460; Patched; 1.
 DR PROSITE; PS0156; SSD; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 974 AA; 112088 MW; 0286DD1D4A6A8892 CRC64;

Query Match 6.7%; Score 464; DB 2; Length 974;
 Best Local Similarity 19.7%; Pred. No. 1.1e-23;
 Matches 221; Conservative 173; Mismatches 386; Indels 344; Gaps 38;

QY 324 SLSDKLS-----FSTHTLGGQFGMGWTWASWPLTILVLSVIVVVALAAGLVF---TELT 376
 Db 7 TLDKLEKFEFAWYTRHLLVDYVY---VFLIAPILTAICSGGFGYWKELT 53
 QY 377 T-----DPV-----ELWSPNSQARSEKAFHQ-----HFGPPFTNQV 410
 Db 54 LLDARKLYTPVSAPSWSREVEFSELPVKSYPFLPERTFQWNRVLYLVVHGRPFENGSY 113
 QY 411 IL-----TAPNRSSYRYDSLLGPKNFSGILDLDLLELLELLELLELLELLEL 455
 Db 114 NLEGSYLLDDIAKIBEEVATNVSFRREALEDNPALM-----RGRHLN 156
 QY 456 VNSPRAQRNISODICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLPANOTLMGQTSQ 515
 Db 157 STEIEVNQNTTI-----TFKDVCMNVYGDYREKNVIELKRRHELN----- 198
 QY 516 VDMKHFLYCANAPLTFKGTALALSCMADYCAPVPPFLAIGG---YKGDVSEAEALIM 572
 Db 199 -----KRGISVTTPQVQNEGTPIYAFVIGGVDTYPNDTIKYARAMEL 241
 QY 573 TFLANNYPAGDPRLAQAKLWE---BAFLEEMRAFORRMAGMFQVTFATERSLEDEINRTT 629
 Db 242 WYFLAKFDEEQQL--AKWEDTAKEFVRE--TYEDHPT--IQCHKHSRIVDQGLTRNA 295
 QY 630 AEDLPFATSYIVIFLYISLALGYS-----SNRVMVDSKATLGLGVA 674
 Db 296 NRLKEVFNVTIAVLFTAF-----YSVKMYFRTHCWPLHIDWLR-----SKPMLALGGVL 347
 QY 675 VVLGAVMAAMGPFSLGIRSSLVILQVVPFLVLSVGADNIPFVLEY-----QRLPRPGE 730
 Db 348 SSVTAILSGIGLLWFGMFFABITL-IAPFLVLSIGVDDMFIAVAAMHNTMKYGRSPK 406
 QY 731 PREVHIGALGRVAPSMLLCSLSAICPFLGALTMPAVRTALTSLGLAVILDFLLQMSA 790
 Db 407 VMKKRMVAMESAVAIPTSTFVLVSGVGIITDIIAVQGCAMTAACMFFFLQITF 466
 QY 791 FVALLSLDSKQESRLDVCCV-----KPOE----- 817
 Db 467 PAALMVISAKAOMSGRNSCMPCTADDIYTIEDGSLQPNLKKSKSRKDAKEKPEKMN 526
 QY 818 -----LPPGQ-----GEGLLGFQKAYAPFLHWTGRVLLFLAL 856
 Db 527 EIAESEIAEKRSYSSDPDSQIHIPVKSRGAMGHFRDYVFWLLNWKTKLFLVFIYV 586
 QY 857 FGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYEFGAPVFTVTLGYNFSEAGM 916
 Db 587 LGTALIGICVMEQGLDYDKLLHSDPLVE---ALKREIDL----- 623
 QY 917 NAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWDDFDIDLWTPSSCCCLYISGPNK 976
 Db 624 -----FHGGDQIEIA----- 633
 QY 977 DKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLPHFLNDRPNKPKPGG-----IA 1030
 Db 634 -----IQNCPNLTIPENRDIEM-----IAQFENISYSLGKGSFWMR 673
 QY 1031 AVYSTSVNLTS-----DQGVLA---SRFAYHKLKNSODYTEALRA---ARLAA- 1075
 Db 674 EYKYSNLTSYLNDRNSWIVGVWSQLFAFYKLM---SQDFVWANEADYDTLEKLSYR 731
 QY 1076 -----ITADLRKVPGTDPAFVFPVYITNVFYEQYLTLPEGLFMLSIC 1119
 Db 732 FRIGVHRLSTPTDLVLIIEELRGVADRPDLDDIVYQQSRAADQLNVLSTIINDTLA 791
 QY 1120 LVPTPAVSCLLGLDLRSGLNLLSIVMLVDVTPGMAWDISYNAVSLINLSVAGMSV 1179
 Db 792 MFCMFCVALIPIPNPI-CALFITLANVTIDIGVIGFLSLWSVKLDPISMITIIMSGFSI 850
 QY 1180 EFVSHITRSPALS-TKPTWLERAKEATISMGSAFAGVAMTNLPGLVLGLAKAQLQIIF 1238
 Db 851 EFSAHITGFPVSNENNLAFRCVDMEKLAWPVVHG--SLSTILGVFLAFIDSVMVLVF 909
 QY 1239 FFRLLNLTLLGLLHGLVFLPVILSYVGVDPVNPALALEQKRAE 1282

336 LAGQFQCGTGWASWPLTILVSLVPVVALAGLV-FTLE-TTDPVELWSAPNSQARSE 393
9 LVNAPFWGPVVRHWRWCFISPLFLTLACSQVGFRRMTLRLVDDPSYFTTSPDARWR 68
394 KAFDQHFQGFRTNQVILTAAPNRSSRYRDSLLGPKNFSGILDLDLLELLEL-QLERL 452
69 ISVFENW-PL--DENKFLPGKSFPAKRFVNIILIRAKDGGSIWQNDVLEIHLAQWIMW 125
453 HLOVSPQAQNISLQDICVAPLNPNTSLYD--CCINSLLQYFQNNRTLLLTANTQTLW 510
126 NISIPTDLLKFNLTQDCLLS-----YDWVCGANEHIQWLLLRNDV-----NQILD 171
511 GQTSQVDWQHELYCANAAA--PLTFKQGTALALSCWADYG-APVFPFLAIGGKGDYSE 566
172 LHPFRGGTKDTPVYLGIGFDGVQFQNGT-----LSDAKLTQLCFYFLKQDQKWBEYS 225
567 ABALINTFSLNNVPAGDPRLAQAKLWEAPLEMRAPFORRMAGMFOVFTTAERSLEIN 626
226 K-----FSYAL-----EYFLNQVVS-----SDVITLSFAHYQSLEDGLD 259
627 RTAEDLPFATSYIVIFLYI---SLALGSYSWSRVMVDSKATLGLGVAVVLGAVMAA 683
260 ENAKAPVPNVFVFLVAMVALVSSFTLKSSATKIDWISSKPLAAAGMFWSTVLSIISA 319
684 MGPPSLGIRSSLVILQVVPFLV-----SVGADNIFIFVLEYQRLPRRP 728
320 FGFLTLGVRYN-VINTIIFPIIGEFQSFQKMEKNQIIPAIGTDDMFLMNACWDQTSKL 378
729 GPREVHIGRALGRVAPSMLLCSLSEACFFGALTPMPAVRTFALTSLGLAVILOPFLQM 788
379 SVPE--RMSKTLSHAGVAVTIITVDVMSFAIGCIITDLFGIOPFCIACVSAFSTYQL 436
789 SAFVALLSL-----DSKRQE----- 803
437 TFFSGAMATGVEVERKCHLFFRTFQLVDISKNERADSKLQIKRSASPAPNYLS 496
804 -----ASRLVCC--CVKQELPPG-----QGEGLLGFQKAYAPPL 840
497 SSNSFSDSDSFSSKKTIPABFAWKQOQSPNSLSKSKSDEEKDRIVHFICKIYGPI 556
841 LHWITRGVVLFLALFGVLSYSMCHISVGLQDELALPKDSYLLDYFLPLNRYFVGAPV 900
557 LNSVRIIPSGLLIFVYLAITAMTCTNFRGLNPGNLVNDHVIYAKIYFSDIKHFWRIGQL 616
901 YFVITLGVNFSSEAGNNAICSSAGCNFNSFTOKIQVATEFPFQSYLAIPASSWVDVFDW 960
617 HVALNPPNLTISENREL-----LKVSAPENTQY-TLREGTVFPLEY 661
961 LTPSCCRILYISGPNKDKPCSTVNSLNCNMSITMGSVRPSVEQFKYLPWFLNDRP 1020
662 LN-----YLSLNAE-----VEDTERLWTKLNSMLKY----- 689
1021 NIKCPGGLAAYSTVNL-T-SDQVLAASRPMAYHKLKNSQDYTEALRAARELAN----- 1075
690 -----TGGSTOWASNLKINKTQSGFQAFQIALKNFVEPNDRKHAQALLRDIADHPFN 744
1076 -----ITADLRKVPCTDPAFVFPVTINV-----FVEQYLTILP-----EGLFMLSCL 1120
745 VVVYHEVSFGNKIL-NDFISHSQCYAQKNIKPLAPPADQYLIILPATIQNVVLSLGM 803
1121 VFTFAVSCLLGLDLSGLNLLSIVMLVDTVGFMAWDISYNVAVSLNLSVAGMSVE 1180
804 ----AVVSPLLVPSLPSGPVIFVSIIVSINIGVGYMTLWGVNLDVAVSMISIMSGFVD 859
1181 FVSHITRSPATKPTWLERAKEATISMGSAVPAGVAMTNLPGLIIVLGLAKQLIQIFPF 1240
860 LSAHIYIAVFTVSHGDT-KORVIGALETLGWPIFOGASST-IAGISILYTVDAIIL-LVFF 916
1241 RLNLITLTLGLHGLVFLPVLIS 1263
917 KTIWLTMLIGATHGLFFIPIFUS 939

RESULT 95

Q6AW15 CABEL
ID Q6AW15 CABEL PRELIMINARY; PRT: 1358 AA.
AC Q6AW15;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Patched family protein 3, isoform a.
GN Name-ptc-3; ORFNames=Y110A2AL.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "genome sequence of the nematode C. elegans: a platform for investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AC024746; AAT92058.1; -; Genomic DNA.
DR Ensembl; Y110A2AL.8; Caenorhabditis elegans.
DR WormPep; Y110A2AL.8a; CE37163.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR000731; SSD 5TH.
DR Pfam; PF02460; Patched; 1.
DR PROSITE; PS0156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1358 AA; 152075 MW; 1E920618B75DF741 CRC64;
Query Match 6.6%; Score 455; DB 2; Length 1358;
Best Local Similarity 22.8%; Pred. No. 7.7e-23;
Matches 215; Conservative 146; Mismatches 373; Indels 209; Gaps 37;
QY 473 APLNPDTSLYDCCI---NSLIQYFQNNRTLLLTANTQTLMGQTSQVDWKDHFYLCANAP 529
DB 449 AKKNPEKAKOLLCLLEYGSSLLKWNQENPERL-----GEFLTKEEP 490
QY 530 LTFKQGTALALSCWADYGAPVFP---LAIGYK---GKDYSEARALINTFSLNNYPAGD 583
DB 491 DYPNGYDVMYTGCKG-FGKKIWEWPEDLIIGGIQDNGKLV-ABALQSVFLVSGAYDVF 548
QY 584 PRLAQAKLWEAPLEMRAPFORRMAG-----MPQVTFTAERSLEINRTTADLPFATS 639
DB 549 ARIKNDKTDSPGLDR-HHFQPMWAGEIISTWQNRFT-KRLYSHELNRERRRQHPHPLASTS 606
QY 640 -----YIVIFLYISLALGSYSWSRVMVDSKATLGLGVAVVLGAV 680
DB 607 IADMLEEFSQFNIIIVIGYILMWIYAAPTQGRFGW-WLAVQSNVALAICGVILVTISS 665
QY 681 MAAGFFSYLGIIRSSLVILQVVPFLVSGADNIFIFVLEYQRLPRRPGEPREVHIGRAL 740
DB 666 ICGLGFATHLGIFNFAATTVVPPFLSLGLGIDDMELLHNYDEIINCNKN---EIGVLL 722
QY 741 GRVAPSMLLCSLSEACFFLGLATPMPAVRTPALTSLGLAVILDFLLQWSAFVALLSLDSK 800
DB 723 KETGMSVMLTSTNNITLAFISGYVLPALRSPSCQTALLAFNLFLFPIPPAMIGIDLR 782
QY 801 RQEAERLDVCCVK-----POELP-----PPQGEGLL 828
DB 793 KQKGRDLAYCSRGNPGOMATFSQSVPSNVSNMSSRAELAGYKQADEYKRHEFWTVG-- 840
QY 829 LGFFOKAYAPFLHWTIRGVLLLFALFGVLSYSMCHISVGLQDELALPKDSYLLDYFL 888
DB 841 -GFLNKIYIPALKNNVVKACVLGTTTAVVFLCYMTSTLGLADVLPHPTPAFLR 899
QY 889 FLNRYFVGAPVYFV-----TTLGVNFSSEAGNNAICSSAGCNFSTQKIQA 937
DB 900 AREQYFSF-YPMFAVLRGDKLDIPNQOQLIBEYRAQLG-----SSKFMK 943

QY 938 TEPPEQSLAIPASSVDDFDLW-----TPSS- 965
Db 944 AEGKLQPY-----WMSMLRVLQSLDMALEKDLAAGKFDLTNGNPIKNGEXKSPESM 996
QY 966 -CCRLYIS-GPNKDFCSTVNSLNCNKMISITWGSVRRPSVEOPHKYLP-WFLNDRPNI 1022
Db 997 IARLVCSFGTNYN--CDGRGKMKWNEV-----INP--EGFYNLTGWFNVD--NM 1044
QY 1023 KCPKGGLAAYSTSVNLTSDDGVLASRSMAYHKPLKNSQ-----DYTBALRAARELA 1073
Db 1045 MYVVSQASFYPTPPGWEYNEKL--AKVVPAAEPLLYSQMPFQNDLIDTPAIVKMBEIR 1102
QY 1074 ANITADLRKVPCTDPAFEVFTYTNVFEQYLITLPGELFMSLCLVPTFAVSCLLGL 1133
Db 1103 ATCE-----BYSEGLSNHPSGIAFTWEQYLT-LRWNLFO-AICII-ALAVFCVISIL 1153
QY 1134 DLRSGLNLSIVMLT--VDTVGFMALWDISYNAVSLNLVSAVMSVEFVSHITRSPA 1191
Db 1154 MNPWAAATLMICIVVITIELGGFMGLMKWNPISAVTLCNAGIGVEFTAHVELAF-L 1212
QY 1192 STKPTWLBRAKEATISMSAVPAGVAMTNLPGLVLGLAKAQLIQIFPRLNLLITLGL 1251
Db 1213 TALGTIDQRLSCLOHMFVPVYHGAISTFL-GVVMVSEDFVVTYFTYMTLLVALGV 1271
QY 1252 LHGLVPLVILSYVCPD-----VNPALALEQKRAEA 1283
Db 1272 FNGLCVLPVILTVGPKPELPTDGS SVLPPPPPLRQOYAEKS 1314

RESULT 96

ID Q6AW16 CAEBL PRELIMINARY; PRT; 1361 AA.
AC Q6AW16;
DT 25-OCT-2004 (TreeBLrel. 28, Created)
DT 25-OCT-2004 (TreeBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TreeBLrel. 28, Last annotation update)
DE Patched family protein 3, isoform b.
GN Name=ptc-3; ORFNames=Y110A2AL.8;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; AC024746; AT92059.1; -; Genomic DNA.
DR Ensembl; Y110A2AL.8; Caenorhabditis elegans.
DR WormBase; WBGene00004210; ptc-3.
DR WormPep; Y110A2AL.8b; C373154.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0008158; F:hedgehog receptor activity; IEA.
DR InterPro; IPR003392; Patched.
DR InterPro; IPR00731; SSD 5TM.
DR Pfam; PF02460; Patched; I.
DR PROSITE; PS50156; SSD; 1.
KW Complete proteome.
SQ SEQUENCE 1361 AA; 152418 MW; C60CC426CDEEF869 CRC64;

Query Match 6.6%; Score 453.5; DB 2; Length 1361;
Best Local Similarity 22.7%; Pred. No. 9.9e-23;
Matches 215; Conservative 146; Mismatches 373; Indels 212; Gaps 37;
QY 473 APLNPNTSLYDCCI---NSLLQYFQNNRTLLLTANQTMQTSQVDWKDHFLVCANAP 529
Db 449 AKKNPEKAKDLLCLEYGSLLKMWQENPERL-----GEFLTKEEMP 490
QY 530 LTFKQGTALALSCWADYGAIPVFPF---LAIGGVK---GKDYSEAEALINTFSLNNYPAGD 583

Db 491 DYPNYGDMVTGGCKG-PGKKIMEWPEDLIIGGIQDNGKLV-AEALQSFLVSGAYDVF 548
QY 584 PRLAQAQKWEBAFLBENRAFQRMAG-----MFQVTFTAERSLEDEINRTTAEIDLPIFATS 639
Db 549 ARIKNDKTDGHPGLDR-HHFQPMWAGEIISTWRNFT-KRLYSHELNRERQHFPLASTS 606
QY 640 -----YIVFLYISLALGSYSSMSRVMDSKATLGLGGVAVLCAV 680
Db 607 IADMLEFSPNYIIIVIGILMVIYAFTQGRQGM-WLAVQSNVALAICGVILVTISS 665
QY 681 MAAMGFPSYLGIRSSVLQVVPFLVLSVGADNIFIIVLEYQRLPRPGEPREVHIGRAL 740
Db 666 ICGLGPATHLGINFAATQVVPFLSLGLGIDDMFLLHNYDEIINCNK-----EIGVLL 722
QY 741 GRVAPSMLLSLSRAICFFIGALTPTMPAVRTFALTSLGLAVLDFLLOMSAFVALLSDSK 800
Db 723 KETGSMVLTISNNILAFISGYVLPIPALRSFCQTAILAFNLIFLFIIPFAMIGIDL 782
QY 801 RQEARSLDVCCKVK-----POELP-----PPGQGE 825
Db 783 RQRKGKDLAYCSRGNFQMATSVPSNVSNNTQMSRAELAGYEKQADEYKHEPWTTV 842
QY 826 GLLLGFFQKAYAPFLHWTIRGVVLLLELALFGVLSYSMCHISVGLDQELALPKDSYLLD 885
Db 843 G---GFLNKIYIPALKNNVVKACVLIGTTTAVVFLGYMTSTLGLLELADVLPHTPPAA 899
QY 886 YFLFLNRYFVGAPVYF-----TTLGNFSEAGNAICSSAGCNCNFTOKI 934
Db 900 FLRAEQVFSP-YPMFAVLRGDKLDIPNQQLIEYRAQLG-----SSKF 943
QY 935 QYATEPPEQSLAIPASSVDDFDLW-----TP 963
Db 944 MIKAEGLQPY-----WMSMLRVLQSLDMALEKDLAAGKFDLTNGNPIKNGEXKSP 996
QY 964 SS--CCRLYIS-GPNKDFCSTVNSLNCNKMISITWGSVRRPSVEOPHKYLP-WFLNDR 1019
Db 997 ESMIAARLVCSFGTNYN--CDGRGKMKWNEV-----INP--EGFYNLTGWFNVD- 1045
QY 1020 PNICKPGGLAAYSTSVNLTSDDGVLASRSMAYHKPLKNSQ-----DYTBALRAAR 1070
Db 1046 -NMVYVSQASFYPTPPGWEYNEKL--AKVVPAAEPLLYSQMPFQNDLIDTPAIVKMBE 1102
QY 1071 ELAANITADLRKVPCTDPAFEVFTYTNVFEQYLITLPGELFMSLCLVPTFAVSCLL 1130
Db 1103 EIRATCE-----BYSEGLSNHPSGIAFTWEQYLT-LRWNLFO-AICII-ALAVFCVI 1153
QY 1131 LGLDLRSGLNLSIVMLT--VDTVGFMALWDISYNAVSLNLVSAVMSVEFVSHITRS 1188
Db 1154 SILMFPWAAATLMICIVVITIELGGFMGLMKWNPISAVTLCNAGIGVEFTAHVELA 1213
QY 1189 PAISTKPTWLBRAKEATISMSAVPAGVAMTNLPGLVLGLAKAQLIQIFPRLNLLITL 1248
Db 1214 F-LTALGTIDQRLSCLOHMFVPVYHGAISTFL-GVVMVSEDFVVTYFTYMTLLVA 1271
QY 1249 LGLLHGLVPLVILSYVCPD-----VNPALALEQKRAEA 1283
Db 1272 LGVFNGLCVLPVILTVGPKPELPTDGS SVLPPPPPLRQOYAEKS 1317

RESULT 97

ID PTR9 CAEBL
AC Q03602;
DT 01-FEB-1994 (Rel. 28, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Patched related family protein 9.
GN Name=ptr-9; ORFNames=F54G8.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;


```
Qy 808 DV-CCCV-----XPQELPPP-----CQEGEL----- 827
Db 535 DVLCCCLPALPSTTHDLNLTORSYHYHOQPOQOQOQNNIEAQRHEETLIECQSKSC 594
Qy 828 ----LLGPFQKAYAPFLHMITRGVW----LLLFLALFGVSLYSMCHISVGLDQELALPK 879
Db 595 LSFSLSRFAIKHYAPF-----ITRSVKVFGMLLSGILGVSLSFASKMLPDGLELTDLPVQ 650
Qy 880 DSYLLDYFLNRYFEVGAPVY-----FVTTLGYNFSSEAGMNAICSSAGCANNFSF--TQ 932
Db 651 NTN-----EHRFLSVQGLFGFYSMFAVTQG-----DFEYENNQ 684
Qy 933 KI--QYATEFPEQSYL-----AIPASSWDDFDMLTPSSCC-----RLYI 971
Db 685 KLLHEYHEAFVRVSHVIXKNDNGGLP-DFWLSLFRDVLNQLRAFDRDYREGRITQERWYS 743
Qy 972 SGPNK-----DKFCPSTVNSLNLCKNCKMSITWGSVRPSVEQFHKYL- 1012
Db 744 NASNDAILAYKLLVQTGHVDNPIDKSLVTQVRLVD-----SEGVINP--EAFYNVLS 793
Qy 1013 PWFNLNDRPNIKCPKGLAAYSTSVNLTSQGQVLASRFMAYHKPLKNSODYTEALRAAREL 1072
Db 794 AWAMND-----VLAYG-----ASQLPFYLHGLSDTADIKTMISQIREL 831
Qy 1073 AANITADLRKVPCTDPAFEVEPYTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLG 1132
Db 832 CSRF--ESRGLPN-----YPSGIPPIFWEQYMLRGLLKAIGCALLAVFVVSLLT- 881
Qy 1133 LDLRSGLLNLLSIVMILVDTVGFMAWDIISYNAVSLINLVSAGVMSVEFVSHI 1185
Db 882 LSGWAALLIVNLVNMWQIQLLGMILLGKLSAIPAVILIASIGLGVGIVTHV 934
```

Search completed: April 11, 2006, 00:45:18
Job time : 316 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:45:32 ; Search time 53 Seconds
(without alignments)
2077.811 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAI SNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/aaa/5 COMB.pep.*

2: /cgn2_6/ptodata/1/aaa/6 COMB.pep.*

3: /cgn2_6/ptodata/1/aaa/H COMB.pep.*

4: /cgn2_6/ptodata/1/aaa/PCTUS COMB.pep.*

5: /cgn2_6/ptodata/1/aaa/RE COMB.pep.*

6: /cgn2_6/ptodata/1/aaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2403	34.8	1318	2	US-09-949-016-10152
2	2402.5	34.8	1278	2	US-09-462-136-2
3	2385	34.5	1319	2	US-09-462-136-4
4	1329	19.2	1170	2	US-09-462-136-6
5	1078	15.6	1296	2	US-09-462-136-9
6	651.5	9.4	1447	1	US-08-540-406-19
7	651.5	9.4	1447	2	US-08-656-055-19
8	651.5	9.4	1447	2	US-08-954-668-19
9	651.5	9.4	1447	2	US-08-954-668-19
10	651.5	9.4	1447	2	US-08-918-658-19
11	651.5	9.4	1447	2	US-09-724-631-19
12	651.5	9.4	1447	2	US-08-954-701A-19
13	651.5	9.4	1447	2	US-09-807-007-6
14	651.5	9.4	1447	2	US-09-754-032-19
15	651.5	9.4	1447	2	US-08-916-140-19
16	651.5	9.4	1447	4	PCT-US95-13233-19
17	646.5	9.4	1434	1	US-08-540-406-10
18	646.5	9.4	1434	2	US-08-656-055-10
19	646.5	9.4	1434	2	US-08-954-668-10
20	646.5	9.4	1434	2	US-08-918-658-10
21	646.5	9.4	1434	2	US-09-724-631-10
22	646.5	9.4	1434	2	US-08-954-701A-10
23	646.5	9.4	1434	2	US-09-754-032-10
24	646.5	9.4	1434	2	US-08-916-140-10
25	646.5	9.4	1434	4	PCT-US95-13233-10
26	639.5	9.3	642	2	US-09-248-796A-16161
27	628	9.1	1296	2	US-08-857-636-60

28	584	8.5	1203	2	US-09-293-505-2	Sequence 2, Appli
29	584	8.5	1203	2	US-09-060-939A-2	Sequence 2, Appli
30	583.5	8.4	1182	2	US-09-293-505-7	Sequence 7, Appli
31	583.5	8.4	1182	2	US-09-060-939A-7	Sequence 7, Appli
32	581.5	8.4	531	2	US-09-248-796A-16160	Sequence 16160, A
33	580	8.4	1203	2	US-09-207-857-2	Sequence 2, Appli
34	580	8.4	1203	2	US-09-909-280A-2	Sequence 2, Appli
35	579	8.4	1146	2	US-09-807-007-1	Sequence 1, Appli
36	526.5	7.6	1311	1	US-08-540-406-4	Sequence 4, Appli
37	526.5	7.6	1311	1	US-08-656-055-4	Sequence 4, Appli
38	526.5	7.6	1311	2	US-08-954-668-4	Sequence 4, Appli
39	526.5	7.6	1311	2	US-08-918-658-4	Sequence 4, Appli
40	526.5	7.6	1311	2	US-09-724-631-4	Sequence 4, Appli
41	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
42	526.5	7.6	1311	2	US-08-916-140-4	Sequence 4, Appli
43	526.5	7.6	1311	4	PCT-US95-13233-4	Sequence 4, Appli
44	526.5	7.6	1311	4	US-09-268-140-3	Sequence 3, Appli
45	474	6.9	1286	2	US-08-460-900C-62	Sequence 62, Appli
46	464.5	6.7	1299	2	US-08-674-509B-48	Sequence 48, Appli
47	464.5	6.7	1299	2	US-08-954-698-48	Sequence 48, Appli
48	464.5	6.7	1299	2	US-09-639-695-62	Sequence 62, Appli
49	464.5	6.7	1299	2	US-09-448-188-48	Sequence 48, Appli
50	464.5	6.7	1299	2	US-08-954-128-48	Sequence 48, Appli
51	464.5	6.7	1299	2	US-08-954-740-48	Sequence 48, Appli
52	464.5	6.7	1299	2	US-08-540-406-6	Sequence 6, Appli
53	438.5	6.3	1285	1	US-08-656-055-6	Sequence 6, Appli
54	438.5	6.3	1285	2	US-08-954-668-6	Sequence 6, Appli
55	438.5	6.3	1285	2	US-08-918-658-6	Sequence 6, Appli
56	438.5	6.3	1285	2	US-09-724-631-6	Sequence 6, Appli
57	438.5	6.3	1285	2	US-08-954-701A-6	Sequence 6, Appli
58	438.5	6.3	1285	2	US-08-916-140-6	Sequence 6, Appli
59	438.5	6.3	1285	4	PCT-US95-13233-6	Sequence 6, Appli
60	438.5	6.3	1285	2	US-09-500-063-2	Sequence 2, Appli
61	438.5	6.3	529	2	US-09-270-767-31821	Sequence 31821, A
62	254.5	3.7	529	2	US-09-252-991A-17657	Sequence 17657, A
63	151	2.2	232	2	US-09-252-991A-16989	Sequence 16989, A
64	150	2.2	1237	2	US-09-328-352-4877	Sequence 4877, Ap
65	145	2.1	1467	2	US-09-328-352-4877	Sequence 12490, A
66	140.5	2.0	1051	2	US-09-489-039A-10700	Sequence 94, Appl
67	140	2.0	1005	2	US-09-248-796A-16571	Sequence 16571, A
68	138	2.0	826	2	US-10-055-364-42	Sequence 42, Appl
69	137.5	2.0	1040	2	US-08-804-439A-14	Sequence 14, Appl
70	136.5	2.0	1032	2	US-08-720-229-14	Sequence 14, Appl
71	133	1.9	792	2	US-09-252-991A-24533	Sequence 24533, A
72	130.5	1.9	808	2	US-09-902-540-16765	Sequence 16765, A
73	130.5	1.9	808	2	US-09-710-279-2020	Sequence 2020, Ap
74	130.5	1.9	808	2	US-09-134-001C-3105	Sequence 3105, Ap
75	130.5	1.8	1069	2	US-09-252-991A-27976	Sequence 27976, A
76	127.5	1.8	1607	2	US-09-252-991A-28767	Sequence 28767, A
77	126	1.8	801	2	US-09-328-352-8064	Sequence 8064, Ap
78	126	1.8	808	2	US-09-252-991A-21000	Sequence 21000, A
79	125.5	1.8	1123	2	US-09-489-039A-13014	Sequence 13014, A
80	125.5	1.8	789	2	US-09-252-991A-30591	Sequence 30591, A
81	123.5	1.8	1065	2	US-09-902-540-10725	Sequence 10725, A
82	123	1.8	1033	2	US-09-902-540-14015	Sequence 14015, A
83	123	1.8	1049	2	US-09-252-991A-20611	Sequence 20611, A
84	122.5	1.8	1052	2	US-10-104-047-2923	Sequence 2923, Ap
85	120	1.7	863	2	US-09-902-540-11066	Sequence 11066, A
86	118.5	1.7	735	2	US-09-902-540-12414	Sequence 12414, A
87	118.5	1.7	1033	2	US-08-311-731A-363	Sequence 363, App
88	115.5	1.7	914	2	US-09-252-991A-24417	Sequence 24417, A
89	115	1.7	651	2	US-09-603-208A-226	Sequence 226, App
90	115	1.7	810	2	US-09-252-991A-21762	Sequence 21762, A
91	114	1.7	976	2	US-09-902-540-11016	Sequence 11016, A
92	114	1.7	1019	2	US-09-543-681A-8143	Sequence 8143, Ap
93	112.5	1.6	697	2	US-09-540-236-2360	Sequence 2360, Ap
94	112	1.6	517	2	US-09-711-164-443	Sequence 443, App
95	112	1.6	750	2	US-09-489-039A-283	Sequence 283, App
96	112	1.6	1051	2	US-09-489-039A-7216	Sequence 7216, Ap
97	111.5	1.6	506	2		
98	111	1.6	1025	2		
99	111	1.6	1025	2		
100	111	1.6	1048	2		

ALIGNMENTS

```
RESULT 1
US-09-949-016-10152
; Sequence 10152, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRN
; ORGANISM: Human
US-09-949-016-10152

Query Match          34.8%; Score 2403; DB 2; Length 1318;
Best Local Similarity 39.7%; Pred. No. 3.9e-218; Indels 110; Gaps 26;
Matches 525; Conservative 238; Mismatches 451;

QY  2  ABAGLRGMLLWA-----LLRLAQSEPYTHIHPGYCAFYDECG-----KNPELSG 47
DB  32  ASAAALRGHSMTARGALGLLLLLCPAQVFSQ-----SCVWIGEGCIAGVDRKRYNCEYSG 86
QY  48  SLMTLSNVSLCNTPARKITGDBHLLILQKICRLTYGPNQACCSAKQLVSEASLSITK 107
DB  87  -----PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDLNQLPL 131
QY  108  ALLTRCPACSDNFVNLHCHNTCSNQSLFINVTR-----VAOLGAGQLPAVAYEAFYQHS 163
DB  132  QFLSRPCPCFYNNLNLFCBLTSCPSQSQFLANTATEDYDPTNQTNTNKLQYVVGQS 191
QY  164  FAEQSYDSCSRVRVFAAATLAVGTGCGVYGSALCNAQRWLNFGQDGTGNGLAPLDTIT---- 219
DB  192  FANAMYNACRDVEAPSSNDKALGLLCKGDADA-CNATNWEYMFKNKNGQAPFTITPVFS 250
QY  220  ----FHLLPEGQAVGSGIQPLNKGVARCNESQGGDVATCSQDCAASCAPAIAPQ----- 270
DB  251  DRPVH-----GMPEPMNATKGCDSVDEVTAPCSCDCSIVCGPKPQPPPPAPW 300
QY  271  ----ALDSTFVLGMPGSLVLIILCSVEAVVITLL-----VGRPVAPARDKS 314
DB  301  TILGLDAMVIMWITMYAFLLVFFGFAVWCVRKRYFVSEYTPDSDNIAFSV-NASDKG 359
QY  315  KMVDPKKGTSLSKLSFSTHTLLGQFFQGGWGTWASWPLTILVLSVIPVVALAAGLVFTE 374
DB  360  E-----ASCCDPVSAAPGEGCLRLIFTRWGSFCVRNPGCVIFFSLVFTACSSGLVFR 412
QY  375  LTTPDVELMSPNSQARSKAHQHQHFGPFRTNQVILTPAPNRSSVRYDSLLGPK-NFS 433
DB  413  VTTNPVDLMSAPSSQARLEKEYPDQHFPGPFRTQELIIRAPLTDKHIQYFPYPSGADVPG 472
QY  434  GILDLDLLELLELQERLHLQWSPQAQRNISLQDI CVAPLNPDNTSLVDCINSLLOY 493
DB  473  PPLDIQLHLQVLDLQIAIEN--ITASYDNETVTQLQICLAPUSPYNT---NCTILSVLNY 527
QY  494  FQNNRTLLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFKDGFTALSCMADYGAVPVFP 553
DB  528  FQNSHSLVDHKKGDDFF---VYADYHTHFYLCVRAPASLNDTSLLDHPCLTGFGGVPFW 584
```

```
QY  554  LAIGGYKGKDYSEAEALIMTFSLNYPAGDPRLAQAQKLWEEAFLEEMRAFORRMAGMFQV 613
DB  585  LVLGGYDDQNNYNATALVITFPVNNYNDTEKLOQAQAWKEKFINFKVKNYK---PNLTI 641
QY  614  TFTAERSLEDEINRTAEDLPPIFATSYIVIFLYTISLALGSYSSWSRWSKATILGGGV 673
DB  642  SFTAERSIEDELNRESDSVFTVVVISYAIMPLYSIALGHKIKSCRRLLVDSKVSLGIAGI 701
QY  674  AVVLGAVMAAGPSPSYLGIRSSVLIOVPLVLSVGADNIFIVLEVQRLPRPGEPRE 733
DB  702  LIVLSSVACSLGVFSYIGLPLTIVIEVIFLVAVGVNDNIFILVQAYQORERLQGETLD 761
QY  734  VHIGRALGRVAPSMMLCSLSEAICFFLGCALTMPAVRTFALTSGLAVIDLDFLLQMSAFV 793
DB  762  QQLGRVLGEVAPSMFLSSFSETVAFFLGALSVMAVHTFSLFAGLAVFIDFLQLITCFVS 821
QY  794  LLSLDSKRQREASRLDVCCCKVPQBLPPGQ-GEGLLLGFPQKAYAPFLLHWHITRGVLL 852
DB  822  LLGLDIKRQEKRLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLLKDMMRPIVIAI 881
QY  853  FLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTILGYN 912
DB  882  FVGLSFSIAVLNKLVDIGLDQSLSPDDSYNVDYFKSISQYLHAGPPVYFVLEEGHDYIS 941
QY  913  EAGNAICSSAGCNKNSFTQKIQVATEFPPEQSYLAIPASSWVDDPFDWLTTP-SSCCRL 971
DB  942  SKGNMVCGGMGCCNDSLVQQIFNAQQLDNVTRIGFAPSSWIDDYFDWVKPQSSCCRV-- 999
QY  972  SGPNKDKPCPSTVNSLNLCKNCSIT-MGSVVRPSVEQPHKYLPMFLNDRPNIKCPKGGL 1030
DB  1000  -DNITDQFCNASVVDPAQVR-CRPLTPBKGQRPQGGDFMRFLPMFLSDNPNPKCGKGHA 1057
QY  1031  AYSTSVN--LTSQDVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGETD 1088
DB  1058  AYSSAVNILLGHGRVGTAGTYFTYHTVLTQTSADFIDALKARLLASNTV-ETMGINGS-- 1114
QY  1089  AFEVFPYFITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLASGLNLISYMI 1148
DB  1115  AYRFPFSVYFVFEQYLTIDDTIFNLGVSLGAIFLVTWVLGCGELWSAVIMCATIAMV 1174
QY  1149  LVDTVGFMALMDISYNAVSLNLSAVGMSVEFVSHITRSFAISKTPTWLERAKATISM 1208
DB  1175  LVNMPGVNWLWGISLNAVSLNLSVMSCGISVEFCSHITRAFTVSMKGSRVERAEALAHM 1234
QY  1209  GSAPVAGVAMTNLPGILVLGLAKAQLIOIFFRFLNLLITLLGLLHGLVFLPVLVSGPD 1268
DB  1235  GSSVFSGITLTKFGGIVVLAFARSKQIFQIFFRYFMYLAWVLGANTHGLIFLPLVLLSY 1294
QY  1269  VNPA 1272
DB  1295  VNKA 1298

RESULT 2
US-09-462-136-2
; Sequence 2, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Caratea, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRN
```



```

; ORGANISM: Homo sapiens
; US-09-462-136-2

Query Match      34.8%; Score 2402.5; DB 2; Length 1278;
Best Local Similarity 39.8%; Pred. No. 4.2e-218;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLRLAQSEPTTHIQPGYCAFYDRCG-----KNPELGGSLMTLSNVCSLS 59
DB 4 RGLALGGLLLLLCPAQVFSQ-----SCWYGCAGIAYGDRYNCBSYG-----46

QY 60 NTPARKITGDHLLILQKICPRLYTGTNTQACCSAKQLVSLASLSTTKALLTRCPACSDN 119
DB 47 --PPPLPKDGYDLQVELCPGFFG--NVLCCDVRQLQTLKONLQLPQLFSLRCSFCFN 103

QY 120 FVNLCHNTCSNQSLF INVT- ---VAQLGAGQLPAVVAEAFQHSFAEQSYDSCSRV 175
DB 104 LNLFLCELTCSPRQSOFLNVTATEDYDVPVTNQTINKVELQYVYVQSPFANNAYNACRDV 163

QY 176 RVPAAATLAVGTWCGYVSGALCNAORWLNFGQDGTNGLAPLDT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA--CNATNWIEYFNKONGQAPFTITPVFSDFFVH-----215

QY 229 VSGGIQPLNEGVARCNESQGDVATCSODCAASCPAIAPO-----ALDSTFYLG 279
DB 216 ---GMEPMNATKGDSDVETAPCSQDCSIVCGPKPQPPPPPPAPMTILGLDAMYIM 272

QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVPGAFVAVWCYKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324

QY 327 DKLSSTHTLLGQFGTGWASWPLTLVLVSIPVVALAAGLVFTLTTPDVLWAP 386
DB 325 DPVSAFEGCLRLRTRWGSFCVPGPCVIPPVSLVFITACSSGLVFRVVTNPVFLWAP 384

QY 387 NSQARSEKAFDHPGPPRTNQVILTAPNRSYRYSLLGPK--NFGSILDLLELL 445
DB 385 SQQALEKEYFQHPGPPRTTQILRAPLTDKHIYQYPVPGADVPFGPPLDIQLHVL 444

QY 446 ELQERLRHLQWSPQAORNI SLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRLLTLTA 505
DB 445 DLQIAIEN--ITASYDNETVTLQDLCLAPLSPYNT---NCTILSVNLFQNSHVDHKK 499

QY 506 NOTLNGQTSQVMDKHDFLYCANAPITFKDGTALALSCMADYGAPEVPPFLAIGYKGDYS 565
DB 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHPCLTGFGPVPVPLVLGGYDDQNY 556

QY 566 EBAALIMTFLSNYPAGDRLAOKLWEAFLEEMRAFQRRMAGMQVTFTHAERSLEBRI 625
DB 557 NATALVITFPVNNYNDTEKLOQAQAEKEFTNFVKNYK---PNLTISFTAERSIEDEL 613

QY 626 NRTTAEDLPIPATSVIVIPFLYLSALGSYSWSRVMDSKATLGLGGVAVVLGAYWAAG 685
DB 614 NRESDSVFTVVISVAINFLYLSALGHIKSCRLLVDKSVLSLGTAGLILVSSVACSLG 673

QY 686 PFSYLGIRSSILVQLVFPFLVLSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
DB 674 VFSYIGLPLTLVIEVIFPFLVAVGVDNIFILVQAYQDRERLQGETLQOGLRVLGEVAP 733

QY 746 SMLLCSLEAICFFLGTALTPMPAVTFTALTSGLAVILDFLOMSAFVALLSDSRQAS 805
DB 734 SMFLSSFSFETVAFPLGALSVMVAVHTFSLFAGLAVFIDFLQITCFVSLGLDILKQEN 793

QY 806 RLDDVCCVKPQLPPPG--GEGLLAGFPQKAVAPFLHWHITRGVVLFLALFGVLSYSM 864
DB 784 RLDFICCVRGAGDGTSVQASECLFRFFKNSVSPILLKDWRPVIALFVGLVLSIAVL 853

QY 865 CHISVGLDQELALPKDYSLLDYFLFLNRYFVEGAPVYFTVLIGNFSSAGNAICSSAG 924
DB 854 NKVDIGLQSLMPDSDSYWDYFKISQYLHAGPPVYFVLEBGHDYTSKQNMVCGMG 913

QY 925 CNPFSFTQKIQVATEFPFQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKQKCPST 983
DB 925 CNPFSFTQKIQVATEFPFQSYLAIPASSWVDDFDWLTP--SSCCRLYISGPNKQKCPST 983

```

RESULT 3

```

US-09-462-136-4
; Sequence 4, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Mus sp.
; US-09-462-136-4

```

```

Query Match      34.5%; Score 2385; DB 2; Length 1319;
Best Local Similarity 40.0%; Pred. No. 2e-216;
Matches 524; Conservative 231; Mismatches 440; Indels 116; Gaps 28;

```

```

QY 14 LLLRLAQSEPTTHIQPGYCAFYDRCGKNPELSSLMTLNSVCLSNTPARKITGDHLL 73
DB 53 LLLLLCPAQVFSQ-----SCWYGCAGIAYGDRYNCBSYG-----46 100

QY 74 LQKICPRLYTGTNTQACCSAKQLVSLASLSTTKALLTRCPACSDNVLHCHNTCSNQ 133
DB 101 VQELCPGLFF--DNVSLCCDIQOQLTKSNLQPLQLFSLRCSFCYNLMTLFCELTCSHQ 159

QY 134 SLFNTVTRVAQLGAGQLPA---VVAEAFQHSFAEQSYDSCSRVRVPAATLAVGTC 189
DB 160 SOFLNVTATEDYFDPKTPENTKNVKELEYVYVQSPFANNAYNACRDVPEAPSNKALGLLC 219

QY 190 GYVGSALCNAORWLNFGQDGTNGLAPLDI-----TFHLLPEQAVGSGIQLNEGVARCN 244
DB 220 GRDARA--CNATNWIEYFNKONGQAPFTIIPVPSLSTL-----GMEPMNATKGCN 270

QY 245 ESQGDVATCSQDCQCAASCPAIAPOALDSTFTYLGMPGSLVLIILCSFVAVVTLLVG 304
DB 271 ESDVETGPCSCQDCSIVCG--PKQILQGGGGGG--GLDAMYVIMVTVVAFVFPVFG 327

QY 305 FRVA-----PARDKSNWDPKKGTSLSDKLSSTHTLLGQF 340
DB 328 ALLAVWCHRRRYFVSEYTPIDSNIAFSVNSDKGE-----ASCCDPLGAFDDCLRRM 380

```



```
QY 822 GQEGLLGFFQKAYAPFLHWTITRGVLLFLALFGVSLYSMCHISVGLDQSLALPKDS 881
D 729 -GNEETKESFLKTFYFKMLTQ--KRLIIISAFWFTSLVFLPEIQGLDQTLAVPQDS 784
QY 882 YLLDYLFLNRYFEVGCAPVYVFTTLGYNFSSSEAGNAICSS-AGCNPFSTQKIQYATEF 940
D 785 YLVDYFKDVISFLNVPVYVVK-NLDLTKRQNOQKICGKFTTCERDSLAVLE---Q8 840
QY 941 PRQSYLAIPASSVDDDFIDWLPSS--CCRLYISGNPKDKFCPSTVNSLCLNKNCSITM 998
D 841 RHRSTITEFLANWLDYFELFAPQNDQCCRL---KGTDEVCPSPSPSRCC-ETCFQ--Q 894
QY 999 GSVR-----PSVEQPHKYLWPLNDRPNKCPKGLAAYSTSVNLTSDGQVLAGRFMAY 1052
D 895 GSWNYNMSGFPBGKPMELYSWIN-APSDPCPLGGRAPYSTAL-VYNETSVSASVFRTA 952
QY 1053 HKPLKNSQDYTEALRAARELANITADLRKVPCTDPAPFVPPYTTNVPYEQYLILPRG 1112
D 953 HPLRSQKDFIQAY-----SDGVRISSSPELDMFAYSPPYIFVQYQTLGPILT 1001
QY 1113 LFMLSLCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMAWLDISYNAVSLINLV 1172
D 1002 LKLGSAIILIFISVFL-QNIRSFLALVVTMIIVDIGAMALLGLISNAVSLVNDI 1060
QY 1173 SAVGMSVEFVSHITRSFAI---STRPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGL 1229
D 1061 ICVGLGVPCVHIVRSFTVVPSETKKDANSRVLYSLNTIGESVIGKITLTKFIVGCVLAF 1120
QY 1230 AXAQILQIPFRLNLLITLLGLHLGLVFLPVLSVYG 1266
D 1121 AQSKIPDFYPRMFWTLIIAVALHALLFLPALLSLFG 1157

RESULT 5
US-09-462-136-9
; Sequence 9, Application US/09462136
; Patent No. 6426198
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/09/462,136
; CURRENT FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 1296
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-462-136-9

Query Match 15.6%; Score 1078; DB 2; Length 1296;
Best Local Similarity 24.0%; Pred. No. 2.8e-92;
Matches 321; Conservative 237; Mismatches 492; Indels 288; Gaps 41;

QY 76 KTCPLRYTGPNTQACCSAKQLVSLBASISITKALITRCPACSDNFVNLHCHNTCSPNQSL 135
D 63 BFCPHLLTGDN-KLCCTPSQABGLTKQIAQARHILGRCPSCFDNFAKLWCBCFTCSPNQD 121
QY 136 FINVTRVAQL--GAGQLFAVAYEAF-----YQHS--FABQSYDCSRVRVPAATLAVG 186
D 122 FVSISEMPKIEKKEGFTPEYQPAEAYVNTVEYRLSTDFAEQMFSSCKDVTFGGQPALRV- 180
QY 187 TWCYGVYGCALCNAQWLNFGQDTGNGL-APLDITFHLLEPQAVCGSIQIP-LNEGVARCN 244
D 181 -MC-----TSTPCPLTNWLFYFGQNLDMNIPHTKFLYDPIKTPSPDRSTYNNVFTGCD 236
QY 245 ESQGDVATCSQDCAACPAIARPOALDSTPYLQM-----PGSLVLIILICSVFAVVT 299
```

```
Db 237 KSARVGVMPACSTSEC--NKREYANLIDLDGKTSQTCNVHGIACLINI FVMLAFIGSLAV 294
QY 300 ILLVGF-----RVAPARDKSKWVDPKKGTSLSKLSFSHTLLGQPF 341
D 295 LUCVGFVFTSYDEDTVNLRFQTSGBESPKRNIK-----RTGAWI-----HNFMENNA 342
QY 342 QCWGVTWASWPLTILVLSVIPVVALAAGLVFTTATDTPVELWAPNSQARSKAPHDQHF 401
D 343 RDIGMAGRNPKSHFPICGAVLIFCLPGMIYHKESTNVVDMSSPRSRARQEWVFNAP 402
QY 402 GPFFTNQVILTAPNRSYRYSLLGPKNFGSGLDLDLLELLELLEQLERLRLHQQWSPRA 461
D 403 GRPQYQOITMLL--SHRDFQSSGKLYGP-----VFHKDIFEBLFDILNAIKNISTQSDG 455
QY 462 QRNISLQDICYAPLNPDTNLSLVCINSLLQVFNQNRITLLLTANQT-----L 509
D 456 -RTITLDDVCTVRPMGPG---YDCLIMPTNTYFQGNKEHDMKSNKBTVESDDDAFYF 510
QY 510 MGQTSQVDWKDHFYCANAPLTFKGTALALSCMADYGAPVFPFLAIGCYKGDYSEARA 569
D 511 SSEATTDWMNHMAACIDQMSQK--TKSGLSCMGTYGCPSPAPNM-VFGKSTNHOANS 567
QY 570 LIMTSLNNYPAGDPRLLAQAKLWREAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRTT 629
D 568 IMWTILVQTQ--RTEPEIQKAEKWEKFLKCKEYREKSPKI-FSPMAERSITDEIENDA 624
QY 630 AEDLPFATSYIVILYISLALGSY-----SSWSRVWDKATLGLGGVAVVGLGVMAAM 684
D 625 KDEIVTVVIALAFLIGYVTFSLGRYFVCENQLWS-ILVHSR----- 664
QY 685 GPFYSYLGRSSILVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPBHVHIGRALGRVA 744
D 665 GF----- 666
QY 745 PSMLLCSLSEACFPFLGALTMPAVRTPALTSGLAVIDLDFLQMSAFVALLSLDSKROBA 804
D 667 -----TDLPAIRTPCLVAGLAVLIDVVLHCTIFLALFVMDTQRELN 707
QY 805 SRLDVCCCKVQQLPPQCGEGL-----LLG-----PFQKAYAPFLHWTITRGV 849
D 708 G-----KP-BFFPFYQIKDGLGAYLIGQRATDTFTMQTFPHFQVAFPLHMTTRIT 758
QY 850 LLLFLALFGVSLYSMCHISVGLDQSLALPKDYSILDYFLNRYFEVGCAPVYVFTTLGVN 909
D 759 GIIFIASITVTVILSSKISVGFQDSMAKTESKYSITHTFAYLDKFPDVGPPVFTVDGELD 818
QY 910 FSSEAGMNAICSSAGCANNPSTQKIQYATEPPEQSYLAIPASSWVDDFDIMLT-PSSCCR 968
D 819 WHRPDVQNKFCCTFPGCSDTSGFNIMNYAVGHTEQTYLSGEMVYNIDNVLWISRSKSPCK 878
QY 969 LYISQPNKDKCPSTVNSL-----NCLKNCSITMGSVRPSVEQFHKYLFWFLNDRPNIKC 1024
D 879 VYVHDPN--TFCSTNRNKSALDDKACRTCMDF----- 908
QY 1025 PKGGLAAYSTSVNLTSDGQVLASRPMAVHKPL--KNSODYTEALRAARELANITADLRK 1082
D 909 ---DGRASFDAISFTSRGRIQASQWTFHKLSISNSDDPIKAMDTAMVSRRLRSI-- 964
QY 1083 VPGTDPAFEVPPYTTINVFYEQYLTILPEGLFMLSCLCLVPTFAVSCLLGLDRLSGLNML 1142
D 965 ---DDTAHVAYSKIPPFYEQYSTIMILITQTLPITVVGFGIICVTLGDVKGACAV 1020
QY 1143 LSIWMLVDTVGFMAWLDISYNAVSLINLVSAVMSVEFVSHITRSPAISTKPTWLERAK 1202
D 1021 ICQVS-----NYPHIVS-GILIEFSVNLKGYACSLRQAKDRAE 1060
QY 1203 EATISMGSAVFAGVAMTNLPGILVGLKAKAQLIQIFFRNLNLLITLLGLHLGLVPLVTL 1262
D 1061 STVSGIPIIUSGPVVTWAGSTMPLSGAHLQIITVYFPKFLITITVSAVHALIITLPIIL 1120
QY 1263 SYVGPDVNPALAEOKRAEBAVAAVMVASCNHPFSRVSTADNIYVN-----HSFBSGS 1314
```

Db	1121	AFGSRGHSSETSTNDNDQHDACVLS--PTAESHSINVEEGILNRPSPLLDASHILDPL	1178
Qy	1315	IKGAGAISNFPNNGROF	1332
Db	1179	LKAEGGIDKAI---GRDF	1193
RESULT 6			
US-08-540-406-19			
; Sequence 19, Application US/08540406			
; Patent No. 5837538			
; GENERAL INFORMATION:			
; APPLICANT: SCOTT, MATHEW P			
; APPLICANT: GOODRICH, LISA V			
; APPLICANT: JOHNSON, RONALD L			
; TITLE OF INVENTION: Patched Genes and their Use			
; NUMBER OF SEQUENCES: 19			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert			
; STREET: Four Embarcadero Center, Suite 3400			
; CITY: San Francisco			
; STATE: CA			
; COUNTRY: US			
; ZIP: 94111			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: Patent In Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/540,406			
; FILING DATE: 06-OCT-1995			
; CLASSIFICATION: 435			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Rowland, Bertram I			
; REGISTRATION NUMBER: 20015			
; REFERENCE/DOCKET NUMBER: a60190-1			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 415-781-1989			
; TELEFAX: 415-398-3249			
; INFORMATION FOR SEQ ID NO: 19:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 1447 amino acids			
; TYPE: amino acid			
; STRANDEDNESS: single			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
US-08-540-406-19			
Query Match 9.4%; Score 651.5; DB 1; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVGSGIOPLNEGVARCNESQDDVATCSCDCAASCPAIAAPQALDSTFYLGQMPGS	284
Db	24	PGPAGGRRRTGGLRAAAPRD-----YLRHPSYCDAAFALEGI---65	
Qy	285	LVLIIILCSFVAVVTILLVGRVAPARDKSMVDPK-----KGTSLSDKLSFSTHTLL	337
Db	66	-----SKGKATGRKAPLMRAKFORLLFLKGCYIQKNC	98
Qy	338	GQFQGGTGWASWPLTILVLPVVALAAGLVFTLTTPVELASAPNSQARSEKAPH	397
Db	99	GKF-----LVVGLLIFGAFVAGLKAANLEINVELWVEVGRVSRRELYT	143
Qy	398	DQHFGPPFRTN-QVILTAPNRSSRYSDSLIGPNRFSGILDLDLLELLE--LQERLRHL	454
Db	144	RQKIGEAMFNPLMIQTPKEG-----ANVTTEALLQHLDSALQASRVHV	190
Qy	455	QWSPERQARNISLDQICVAP-----LNPDMTSLYDCINSLLQYFQNNRTLLLLTA	505
Db	191	YMYN---ROWKLEHLCKYSGELITETGYMDQIIIEYLIPCLIIITPDCFWEGAKLQSGTA	246

RESULT 7

US-08-656-055-19
; Sequence 19, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;
QY 225 PQAGVSGIQPLNEGVARCNESQGDVATCSCQDCAACPAIARPOALDSTFYLGQMPGS 284
DB 24 PGRPAGGRRRTTGLRRAAAPDRD-----YLRHPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSVFAVVTLLVGFVRVAPARDKSKWDPK-----KGTSLSDKLSFSFTHLL 337
DB 66 -----SKGKATGRKAPLWRAKFORLLFLKLCYIQKNC 98
QY 338 GQFFQCGMTWASWPLTILVLSVIPWALAAGLVFTELTDPELWSAPNSQARSEKAFH 397
DB 99 GRF-----LVGULLIFGAFAGVLKKAANETNVEELWVEVGGVRSRELNT 143
QY 398 DQHFQGPFFRTN-QVILTAPNRSYRYSLLGPKQKPSGIGLDLLELELE--LQRLRLH 454
DB 144 RQKIGEAAMFNQMIQTPKESG-----ANVLTTEALLQHLDSALQSRVHV 190
QY 455 QWSPFAQRNLSLDICVAP-----LNPNTSLYDCCINSLLQYQNNRTLLLLTA 505
DB 191 YMYN-----RQWKLEHLCKYSGELITETGYMDQIIIEYLPCLITPLDCFWGAKLQSGTA 246
QY 506 NQTLMQQ-----TSQVDWKDHPY-----CAN----- 527
DB 247 ---YLLGKPLRNTWDFLEBELKKINYQVDSWBEMLNKABVGHGYMDRPRCLNPADPDC 304

RESULT 8

US-08-954-668-19
; Sequence 19, Application US/08954668
; Patent No. 6172200

QY 528 -----APL-----TFKROCT-----ALALSCMA 544
DB 305 PATAPNKNSTKPLDMALVNLGCGHLSRKYMHQEBELIVGGTVKNSTGKLVSAHALQTMF 364
QY 545 DYCAPVPPPLAIGGVKGDYSAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRAFQ 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE-----DKAAAILSAWORTYVEVHVQSV 411
QY 605 REMAGMFQVTTAERSLEDEINRTTAEDLPIFATSVIVIFLYISLALGSYSSRSRVWVDS 664
DB 412 AQNSTQKVLSTFT--TLLDILKSPDSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY 665 KATLGLGGVAVVLGAVMAAMGFYSYLGRSSVILQVWPFVLVSVGADNIPFVLEBYQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLALGVGDVDFLLAHAFSET 525
QY 725 PRPGEPREHVHIGRALGRVAPSMILCSLSEATCFPLGALTMPNAVTRFALTSLAVILDF 784
DB 526 GQNKRIFFEDRTGECLEKRTGASVALTSISNVTAFFMAALIPALRAFSLQAIAVVVFN 585
QY 785 LIQMSAFVALLSLDSKROEASRLDVCCC-----VKQDEL-----PPPG 822
DB 586 AMVILIFPAILSMNDLYRREDRLDIFCFTSPCVSRVIOVFQAYTDHNDTRYSPPPY 645
QY 823 QGEGE----- 827
DB 646 SSHSPAHEQTITMQSTVQLRTEYDPHTHVYTTABRSEISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGPFQKAYAPFLHWHITRGVWVLLFLALFGVSL 861
DB 706 SSTDLLSQFSDSLHLEPPCTKWTLSFAEKHYAPFLPKPKAVVVFLLFGLGLGSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLNRFVFGVAPVYFVTTLCYVSSBAGNNAICS 921
DB 766 YGTTVRDGLDLDIVPRETREYDFIAAQFKYFSP-----YNN----- 803
QY 922 SAGCNNFSTQKIQYAT-----BPEQSYLAIPASS-----WVDDFDIML----- 961
DB 804 -----YIVTQKADYPNIQHLLDYLRHSFNKVKYVWMLBENKQLPKMWLHVFRDWLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPKDKFCPTVNSLNCCLKMCSITMGS 1000
DB 858 AFDSDMETGKIMPNNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KORLVADAGI 913
QY 1001 VAPSVVEQPHYL-PMFLND-----RNKCPKGLAAY--STSVNLTSDGQ 1043
DB 914 INPSA--FYIYLTAVVSDNPVAYAAQANIRPHFEWHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASRFMAYHKPLKNSQDYTEALRAARELAANITA-DLEKVFQTDPAPEVFPYTTNVFY 1102
DB 969 IEYAQFPFVYNGLRDTSDFVEAIEKVRTICSNYTSLSGLSSYPNG-----YFP---LFW 1018
QY 1103 EGYLTILPEGLPMLSLCLVPTFPAVSCLLGLDLRSLMLLSIVMILVDTVTFGMALWDIS 1162
DB 1019 EGYILGRHMLLFIISVLACTFLVCVFLNPNWTAGII--VMVIALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINLVAAGVSGVSVSHITRSP--AISTKPTWLERAKEATISMGSAVAGVAMTN 1220
DB 1078 LSAVPVILIASVGVGEFTVHVALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LPGAIVLGLAKAQLQIIPFFRLNLLITLGLLHGLVFLPVLISYVG--PDVNPALALEQ- 1277
DB 1134 LLGLVLMAGSBDFIVRYFFAVLAITLITGLVNLGLVLLPVLISFPFGYPVSPANGNLRL 1193
QY 1278 --KRABEAVAAMVWASCNHPRSRVSTADNIYVNHSPGSIKG 1317
DB 1194 FTPSPPEPPSVVRFAAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

Qy	1001	VRPSVEFHKL-PWFLND-----RPNKCPKGGLAAY--STSVNLTSQDQ	1043
Db	914	INPSA--FYIYLTAWNSDPVAYASQANIRPHRPWVHDK---ADYMPETRLRIPAAEP	968
Qy	1044	VLASREWAYHKPLKNSODYTEALRAARELAANITA-DLRKVPOTDPAFEVPPYITNVFY	1102
Db	969	IEVAQFPFYINGLRTDSFVEAIEKVKRTICSNYTSLGLSSYPNG-----YPF---LFW	1018
Qy	1103	EQYLTILPEGLFMLSCLVPTFAVSCLLIGLDRSLGNNLLSIVMLIVDTVGFMAWDIS	1162
Db	1019	EQYIGLRHMLLFIISVVLACTFLVCAVFLNPNWTAGII-VVVLALMTVELFGMGLIGIK	1077
Qy	1163	YNAVSLINLSVAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVFAVAMTN	1220
Db	1078	LSAVPVVILLASVIGVEFVHVHVALAFLTAIGDN---REAVLALEHMFAPVLDG-AVST	1133
Qy	1221	LPGLVILGAKAQLIQIFFFRLLNLLITLGLLHGLVPLPVILSVG--PDVNPALAEQ-	1277
Db	1134	LLGVMLAGSEPDFIVRYFPFVLAAILTLGLVNLGLVLLPVLLSPFGPYPEVSPANGLNRL	1193
Qy	1278	--KRAEBAVAAMVASCNPHPSRVSTADNIYVNHSPGSIKG	1317
Db	1194	PTSPPEPPSVVRFAAMPBGHTH--SGSDSDSESYSSQTTVSG	1233
RESULT 13			
US-09-807-007-6			
; Sequence 6, Application US/09807007			
; Patent No. 6881833			
; GENERAL INFORMATION:			
; APPLICANT: ZAPHIROPOULOS, Peter et al.			
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALING PATHWAY			
; FILE REFERENCE: 2921-0130P			
; CURRENT APPLICATION NUMBER: US/09/807,007			
; CURRENT FILING DATE: 2001-04-06			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 6			
; LENGTH: 1447			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-807-007-6			
Query Match 9.4%; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVGSGIQLNEGVARCNESQGGDVATCSQDCAASCPAIARPOALDSTFYLGQMPGS	284
Db	24	PGRPAGGRRRTTGLRRAAAPDRD-----YLHRPSYCDAAFALEQI---	65
Qy	285	LVLIIILCSFVAVVTILLVGRFVAPARDKSKWDPK-----KGTSLSDKLSPSTHTLL	337
Db	66	-----SKGATGRKAPLWLRAPQRLFLKGCYIQKNC	98
Qy	338	GQFFQGWGTWASWPLTILVLSVPVVALAAGLVFTETDTPVELWSAPNSQARSEKAPH	397
Db	99	GKF-----LVVGLLIFGAFVGLKAAANLETNVEELWVEVGGVRSRELYNT	143
Qy	398	DQHFQGFPRTN-QVILTAPNRSSRYDLSLLGPKNFSGILDLDLLELLE--LQERLRL	454
Db	144	RQKIGEEAMFNPQLMIQTPKEEG-----ANVLTTEALLQHLDSALQASRVHV	190
Qy	455	QWVSEAPQARNISLQDICYAP-----LNPDTNSLYDCCINSLLQVFNRRITLLTLTA	505
Db	191	YMYN-----RQWKLEHLCYSGELITETGYMDQIIYLPCLITPLDCFWEGAKLSGTA	246
Qy	506	NOTLMGO-----TSQVDMKDHFLY-----CAN-----	527
Db	247	--YLLGKPLRWTNFDPLEFLEELKKNYQVDSWEEMLNKAEBVGHGYMDRPLNADPDC	304
Qy	528	-----APL-----TPKDG-----ALALSCWA	544
Query Match 9.4%; Score 651.5; DB 2; Length 1447;			
Best Local Similarity 22.0%; Pred. No. 1.1e-51;			
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;			
Qy	225	PGQAVGSGIQLNEGVARCNESQGGDVATCSQDCAASCPAIARPOALDSTFYLGQMPGS	284
Db	24	PGRPAGGRRRTTGLRRAAAPDRD-----YLHRPSYCDAAFALEQI---	65
Qy	285	LVLIIILCSFVAVVTILLVGRFVAPARDKSKWDPK-----KGTSLSDKLSPSTHTLL	337
Db	66	-----SKGATGRKAPLWLRAPQRLFLKGCYIQKNC	98
Qy	338	GQFFQGWGTWASWPLTILVLSVPVVALAAGLVFTETDTPVELWSAPNSQARSEKAPH	397
Db	99	GKF-----LVVGLLIFGAFVGLKAAANLETNVEELWVEVGGVRSRELYNT	143
Qy	398	DQHFQGFPRTN-QVILTAPNRSSRYDLSLLGPKNFSGILDLDLLELLE--LQERLRL	454
Db	144	RQKIGEEAMFNPQLMIQTPKEEG-----ANVLTTEALLQHLDSALQASRVHV	190
Qy	455	QWVSEAPQARNISLQDICYAP-----LNPDTNSLYDCCINSLLQVFNRRITLLTLTA	505
Db	191	YMYN-----RQWKLEHLCYSGELITETGYMDQIIYLPCLITPLDCFWEGAKLSGTA	246
Qy	506	NOTLMGO-----TSQVDMKDHFLY-----CAN-----	527
Db	247	--YLLGKPLRWTNFDPLEFLEELKKNYQVDSWEEMLNKAEBVGHGYMDRPLNADPDC	304
Qy	528	-----APL-----TPKDG-----ALALSCWA	544
Db	305	PATAPNKSTKPLDMALVINGCHLSRKMYHQBELIVGGTVRNSTKLVSAHALQTMF	364
Qy	545	DYGAVPEPLAIGVYKGDYSEALIMTFSLNNYPAGDPLAQAQKWEAEAFLEEMRAFQ	604
Db	365	QUMTPQMY---EHFKGYEY-----VSHINWNE-----DKAAAILEAQRYVEVHVOSV	411
Qy	605	RRMAGMFQVTTAERSLEDEINRTAEDLPFATSYVIFLYISIALGYSWSRWVDS	664
Db	412	AQNSTQKVLSTFT--TTLDILKSFSVSVIRVASGYLLMLAYACLTM---LRWD---CSKS	465
Qy	665	KATLGLGVAVLGVAMWAGFFSYLGRSSVLIVQVPLVLSVGDNIPIFVLEYQRL	724
Db	466	QGAAGVLGALLVALSVAAGLGLCSLIGISFNAATTQVLPFLGALGVGDVDFLLAHAFSET	525
Qy	725	PRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGLALTMPAVRTFALTSLGLVILDF	784
Db	526	GQNKRIPFEDRTGECLEKRTGASVALTSISNVTAFWNAALIPIPALRAFSLQAUVVVFNF	585
Qy	785	LLQMSAFVALLSDSKRQASRLDVCCC-----VKQOEL-----PPPG	822
Db	586	AMVLLIFPAILSDMLYRREDRLDIFCCFTSPCVSRVIOVEPQAVTDTHTNTRYSPPPPY	645
Qy	823	QSEGL-----	827
Db	646	SSHSFAHQITWQSTVQLRTEYDPTHVYTTAEPRSEISVQPVTVTQDTLSCQSPST	705
Qy	828	-----LLGFFQKAYAPFLLHMITRGVILLFLALFGVSL	861
Db	706	SSTRDILLSQFSDSLHCLBPPCTKWTLSSFAEKHYAPFLKPKAKVWVIFLGLIGVSL	765
Qy	862	YSMCHISVGLDQELALPKDSYLLDYFLFLNRYVEGAPVYVTTLGYNFSSENGWNAICS	921
Db	766	YGTTRVRDGLDLDIVPRETREYDFIAAQPKYFSF-----YNN-----	803
Qy	922	SAGCNFSFTQKIQYAT-----ERPEQSYLAIPASS-----WVDDFDIDLW----	961
Db	804	-----YIVTOKADYPIQHLLYDLHRSFNSVNYKVMLEENKQLPKWMLHYFRDNLQGLQD	857
Qy	962	-----TPSS-----CCRLYISGNKDKFCPSTVNSLNCNCKMSTIMGS	1000
Db	858	AFDSDMETGKIMPNYKNGSDGVLAVKLLAVQTGSRDK--PIDISQLT--KORLVDADGI	913

Db 305 PATAPNKSTKPLDMALVNLGGCHLSRKYMHWQBELIVGGTVTKNSTGKLVSAHALQTMF 364
QY 545 DYGAVPFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLMEAFQ 604
Db 365 QLMTPKQMT---EHPKGYEY-----VSHINWNE-----DKAAAILLEAWQRTVVVHQS 411
QY 605 RWMAGHPQVTPAERSLEDEINRTAEDLPIFATSVIVIFLISALGSSYSWSRWVDS 664
Db 412 AQNSTQKVLSEFTT-TTLDILKSFSDSVIRVASGYLLMAYACLTM---LRWD---CSKS 465
QY 665 KATGLGVAVVGLVGAAMAGFFSYLGRSSVILQVVPFLVLSGADNIFIVLEYOQL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTVQLFPLALGVGVDDVFLAHAFSET 525
QY 725 PRPGEPREHVHTRALGRVAPSMGLCSLSEACFPFLGALTMPAVRTFALTSLGLAVIDP 784
Db 526 GQMKRIPPEDRTGECKRTGASVALTSISNVTAFFMAALIPALRAFSLQAQVAVVFN 585
QY 785 LLOMSAFVALLSLSKQASRLDVCCC-----VKQEL-----PPPG 822
Db 586 AMVLLIFPAILSMILYRREDRLDFCCFTSPCSRVIQVEPQAYTDDHNTRYSPPPY 645
QY 823 QCEGL-----827
Db 646 SSHSFAHETQITMQSTVQLRBYDEPHVYVYTTAPRSEISVQPVTVTQDTLSCOSPST 705
QY 828 -----LLGFFOKAYAPFLHWHITRGVLLLFALFGVSL 861
Db 706 SSTDRLLSQFSSSLHCLPCKTKWTLSSFAEKHYAPFLPKAKVWVIFLGLLGVSL 765
QY 862 YSMCHISVGLDELALPKDSYLLDPLFNLNRFVEGAPVYVTTTLYGYNPSSAGNAICS 921
Db 766 YGTRVRDGLDLDIVPRETREYDFIAQKYSF-----YNN-----803
QY 922 SAGCNESPTQIOYAT-----EPPEQSYLAIPASS-----WVDDRIDML-----961
Db 804 -----YIVTQADFNQHLLDYLRHSFNKYNVWLBENKQLPKMWLHYFRDMLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDFPCPSTVNSLNLKNCMSITMGS 1000
Db 858 AFSDSWETGKINPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
QY 1001 VRPSVEQFKYL--PWFLLND-----RNIKCPKGLAAY--STSVNLTSDQ 1043
Db 914 INPSA--FYIYLTAWVSDPVAYAAQANIRPREWVHDK---ADYMPETRLRIPAAEP 968
QY 1044 VLASREMAVHKPKNSQDYTEALRAARELANITA-DLRKVPCTDPAPEVPEYTTITNVPY 1102
Db 969 IEYAQFPFYLNGLRDTSDFVEAIEKVRTICSNYTSGLSSYPNG-----YPF-----LFW 1018
QY 1103 EOYLTLPEGLFMLSICLVPTFAVSCLLIGLDLRGLNLNLSIVMLVDTVGFMAWDIS 1162
Db 1019 EGYIGLRHWLLFISVVLACTFLVCVAVFLNFWTAGII-VWVLAJMTVELFCMGLIGIK 1077
QY 1163 YNAVSLINLVSAGMSVRFVSHITSP--AISTKPTWLERAKBATISGMSAVFAGVAMTN 1220
Db 1078 LSAVPWVILIASGVIGVEFTVHALAFLTAIGDKN---RRVLALEHMFAPVLDG-AVST 1133
QY 1221 LPGILVLGAKAQLQIOPFRNLNLTILGLLHGLVFLPVILSYVG--PDVNPALALSO- 1277
Db 1134 LLGLVLMAGSEPDFIVRYFPAVALIITLGLVNLGLVLLPVLSSFFGYPPEVSPANGLNRL 1193
QY 1278 --KRAEAAVAAVWASCPNHRPSRVSTADNIYVNHSHFEKSG 1317
Db 1194 PTPSEPPSVVRPAMPFGHTH--SGSDSDSEYSSQTTVSG 1233

RESULT 14

US-09-754-032-19

; Sequence 19, Application US/09754032

; Patent No. 6921646

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

GOODRICH, LISA V
JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
CITY: Four Embarcadero Center, Suite 3400
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/754,032
FILING DATE: 03-Jan-2001
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/540,406
FILING DATE: 06-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-754-032-19

Query Match 9.4%; Score 651.5; DB 2; Length 1447;

Best Local Similarity 22.0%; Pred. No. 1.1e-51;

Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQAVGSGIQPLNEGVARCNESQGDVATCSQDCAASCPALARPQALDSTFVLGQMPGS 284
Db 24 PGRAGGRRRTTGLRRAAPDRD-----YLRFSYCDAAAPALQI----65
QY 285 LVLIILCSVFAVVITLLVGFVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db 66 -----SKGATGRKAPLWLRAKFORLLFKLGCYIQKC 98
QY 338 QPFGMGWTWASWPLTLLVLSVPVVALAGLVTELTTPDVELWSAPNSQARSEKAPH 397
Db 99 GKF-----LVVGLLIFGAPVGLKAANLETNVEELWVEVGGVRSBELNYT 143
QY 398 DQHEGFPFRTN-QVLTAPNESSRYDLSLLGPKNFSGILDLLLELLE--LQERLRL 454
Db 144 RQKIGEEAMFQPMIQPKKEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QVMSPEAQRNLSLODICYAP-----LNPNDTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 191 YWYN-----ROWKLEHLCYKSGELITETGYMDIIBLYPCLITPDCFWGAKLQSGTA 246
QY 506 NOTLMGO-----TSQVDMKHFLY-----CAN-----527
Db 247 --YLLGKPLRWTFNPDPLFLEELKKNYQVDSWEMLNKAEGVGHYMDRPLCNAPEDPC 304
QY 528 -----APL-----TFKDDT-----ALALSQWA 544
Db 305 PATAPNKSTKPLDMALVNLGGCHLSRKYMHWQBELIVGGTVTKNSTGKLVSAHALQTMF 364
QY 545 DYGAVPFPFLAIGYKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLMEAFQ 604

Db	365	QLMTPKQMY---BHEKGYEY-----VSHINWNE-----DKAAALAEAWQRTYEVVHQSV 411
Qy	605	RRWAGHFQVFTTAERSLEDEINRTTAEDLPIPATSYIVIVFLYISLALGSSYSSKSRVWVDS 664
Db	412	AQNSTOKVLSFTT--TTLDDTLKFSFDSVSRVASGYLLMLAYACTLM-----LRWD---CSKS 465
Qy	665	KATILGIGGVAVLGAWMAAGPFYSILGRSSVILQVWPEPLVLSVCADNIFIVLBYORL 724
Db	466	QGAVGLAGVLLVALSVAAGLUGLCSLIGISFNAATTQVLFFLAGVGVDVFLLAHAFSET 525
Qy	725	PRRPGPREVHIGRALGRVAPSMLLCSLSSEAI CFFL GALT PMPAVRTALTSLAVILDF 784
Db	526	GQNKRPEDRTEGCLKRTGASVALTISINVTAFMAALIPIPALRAFSLQAADVVFNF 585
Qy	785	LLQMSAFVALLSDSKQERASRLDVCCC-----VKPEL-----PPPG 822
Db	586	AMVLLIPAILSMDLVREDRLDIFCCFTSPCVSRVIOPEQAYTDTHNTRYSPPPY 645
Qy	823	QGEGL-----827
Db	646	SSHSPAHEQTIMQSTVOLREYDPTHVYVTTABPRSEISQPVVTQDTLSCQSPST 705
Qy	828	-----LLGFFQKAYAPFLLHMTIRGCVLLLLFLALFGVSL 861
Db	706	SSTRDLLSQSDSLHCLERPCTKWTLSFPAEKHYAPFLKPKAKVVVFLGLILGVSL 765
Qy	862	YSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGPVYFVTTGLGYNFSEAGMAI CS 921
Db	766	YGTRVRDGLDLDIVPRETREYDFIAAQFKYFSF-----YNM-----803
Qy	922	SAGCNNSFTQIOYAT-----EFPEQSYLAIPASS-----WVDDFIOWL-----961
Db	804	-----YIVTQKADYPNIQHLLYDLHRSFNVKYVMLEENKQLPKWMLHYFRDWLQGLQD 857
Qy	962	-----TPSS-----CCRLYISGNKDKFCPSVTNSLNCNKCMSITMS 1000
Db	858	AFDSDWETGKIMPNYKNGSDGVLAYKLLVQTGSRDK--PIDISQUT--KQRLVDADGI 913
Qy	1001	VRPSVEOPHKYL-PWFLLND-----RPNIKCKPKGGIAAY--STSVNLTSDGQ 1043
Db	914	INPSA--FYIYLTAVNSNDPVAYAASQANIRPHPEWVHDK--ADYMPETRLRIPAEEP 968
Qy	1044	VLASRFWAYHKPLKNQODYTEALRAARELANITA-DLRKVGTDPDAFEVPFYTTINVFY 1102
Db	969	IEYAQFPFYINGLRDTSFDEAIEKVRTICSNYTSLSGSSYPNG-----YPF-----LFW 1018
Qy	1103	EQYLTILPEGLFMLSCLVPFPFAYSCVLLGLDLRSGLLNLLSVIMLVDTVGMALWDIS 1162
Db	1019	EQYGLRHLWLLLFISVVLACTFLVCAPVFLNMPWTAGII-VNVLALMTVFLFGMGLIGIK 1077
Qy	1163	YNAYSLINLVAAGMSVEFVSHITRSF--AISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db	1078	LSAPVFWLIIASVGIGVEFTVHVVALAFALGDKN---RRAVLALEHMFAPVLDG-AVST 1133
Qy	1221	LPGLIVLGLAKAQLTIQIPFFRLNLLITLLGLLHGLVFLPILSYVG--PDVNPALALEQ- 1277
Db	1134	LLGVMLAGSEFDFIVRYFFAVLAILITLIGVLINGVLLPVLLSFFGYPVEVSPANGLNRL 1193
Qy	1278	--KQABEAAVAAVMVASCNPHSPRSTADNIYVNHISFECSIKG 1317
Db	1194	PTSPSEPPPSVVRFAMPBGHTH--SGSDSDSEYSSOTTVSG 1233

RESULT 15

RESULT IS
US-08-916-140-19

US-08-916-140-19
: Sequence 19. Application US/08916140

; Sequence 19, Appl. No. 6,946,257
; Patent No. 6,946,257

; PACIFIC NO. 6946237
; GENERAL INFORMATION:

GENERAL INFORMATION:
APPLICANT: Scott Mathew P.

APPLICANT: SCOTT MALINOW F.
APPLICANT: Goodrich, Lisa V.

APPLICANT: GOODRICH, LISA V.
APPLICANT: JOHNSON, RONALD L.

APPLICANT: Epstein, Ervin Jr.


```

Db      646 SSHSFAHETQITMQSTVOLRTEYDPTHVYTTABRSEISVQPVTVTQDTLSQCSPEST 705
QY      828 -----LLGFFOKAYAPFLHMITRGGVLLFLFALFGVSL 861
Db      706 SSTRDLLSQSDSSHLCLPEPCKWTLSFPAEKHYAPFLPKPAKVVFVFLFGLGVSL 765
QY      862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPVYFVTTGLNFSSEAGMNAICS 921
Db      766 YGTRVRDGLDLDIVPRETREYDFIAAQPKYFSF-----YNN----- 803
QY      922 SAGCNFSFTQKIQYAT-----ERPEQSYLAIPASS-----WVDDFDIML----- 961
Db      804 -----YIVTQADYNIQHLLYDLHRSFNSVYVMLEENKQLPKWMLHYFRWLQGLQD 857
QY      962 -----TPSS-----CRLYISQFNKDKFCSTPVSNSLNCIKNCMSITMGS 1000
Db      858 APDSDMETGKIMNNYKNGSDGGLAYKLLIVQTSRDK--PIDISQLT--KQLVDADGI 913
QY      1001 VRPSVQFHKYL--PWFLND-----RPNIKCPKGLAAY--STSVNLTSDDQ 1043
Db      914 INFSA--FYIYTAWVSDNPVAYASQANIRPHRPEWVHDK--ADYMPETRLRIIPAAER 968
QY      1044 VLASRWMAHYKPLKNSQDYTEALRAARELANITA-DLRKVCTDPAPFPPYITINVFY 1102
Db      969 IEYAQPFYGLNGURDTSDFVEAIEKVRTICSNYSIGLSYPNG-----YPP-----LFW 1018
QY      1103 EQYLTLPBGLFMSLCLVPTFAVSCLLGLDLRSGLNLLSVMTLVDTVGPMALWDIS 1162
Db      1019 EQYIGLRHMLLFISSVVLACTFLVCVFLNPNWTAII--VWVLALMTVELFGMGLIGIK 1077
QY      1163 YNAVSLINIVSAGMSVEFVSHITRSP--AISTKPTWLERAKEATISMGSAVPAGVAMTN 1220
Db      1078 LSAVPVILLIASVIGVEFTVHALAFLTAIGDKN--RRAVLALRHMPAPVLDG-AVST 1133
QY      1221 LGCILVGLAKAQLOIOPFRMLNLLITLIGLLHGLVFLPVILSYG--PDVNPALALEQ- 1277
Db      1134 LGLVLMAGSEDFIVRYFPAVALAITLIGVLNGLVLLFVLSFFGYPFVSPANGLNRL 1193
QY      1278 --KRAEEAVALVMAVSCPNHPSRVSTADNIYVNHSPGSIKG 1317
Db      1194 PTPSPPPSVVRFPMPGHTH--SGSDSDSEISSQTTVSG 1233

RESULT 16
PCT-US95-13233-19
; Sequence 19, Application PC/TUS9513233
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCES/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989

```

```

; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13233-19

Query Match      9.4%; Score 651.5; DB 4; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.1e-51;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY      225 PGQAVSGGIQLNBEVGARCNESQGDVATCSCQDCAASCPAIAIPQALDSTFVLGMPGS 284
Db      24  PGRPAGGRRRTTGGLRRAAADPRD-----YLRPSYCDNAFALEQI--- 65
QY      285 LVLIILCSVFAVVTTLLVGVFRVAPARDKSKMVDPK-----KGTSLSDKLSFSTHTLL 337
Db      66  -----SKGKATGRKAPLWLRKAFQRLLPKLGCIYQKNC 98
QY      338 GQPPQCGWTWASWPLTILVSVIPVVALAAGLVFTTELTDTPVELMSAPNSQARSKAFH 397
Db      99  GKF-----LVWGLLIIFGAPAVGLKAANLETNVEELWVEVGGRVSRBLNYT 143
QY      398 DQHFPPFRTN-QVILTAPNRSSRYDSSLGPKNFSGLDLDLLELLE--LQERLRHL 454
Db      144 RQKIGEEAMPNQLMIQTPEBG-----ANVLTTEALLQHLDSALQASRVHV 190
QY      455 QVMSPEAQRNLSLODICYAP-----LNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db      191 YMYN----ROWKLEHLCYKSGELITETGYMDQIIEVLYPCLITPLDCFWEGAKLSGTA 246
QY      506 NOTLWQO-----TSQVDKWKOHLY-----CAN----- 527
Db      247 --YLLGKPLRWTFNDFLEBELKKINYQVDSWEEMLNKAEGYGHGMDRCPCLNPADPDC 304
QY      528 -----APL-----TFKQGT-----ALALSQWA 544
Db      305 PATAPNKNSTKFPLDMALVNGCHGSLRKYKMWQBELIVGGTVKNSTGKLVSAHALQTFP 364
QY      545 DYGAFFVFLAIGYKGYSEAEALIMTFSLNNYPAGDPRLAQAQKLWBEAFLEENRAFQ 604
Db      365 QLMTPKQWY---EHPKGYEY-----VSHINWNE---DKAAALEAWQTYVEVVHQS 411
QY      605 RRMAGHFQVTTAERSLEDEINRTTAEDLPITATSYIVIFLYISIALGYSYWSRWVDS 664
Db      412 AQNSTQKVLSPFT--TTLDLILKSPDSVSVIRVASGYLLMLAYACLTM---LRWD--CSKS 465
QY      665 KATLGLGGVAVVLGAVMAAGPFSYLGIRSSLVILQVVPPLVLSVCAADNIFIVLEYQRL 724
Db      466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTQVLPFLFALGVGVDDVFLAHAFST 525
QY      725 PRPGEPREVIHIGRALGRVAPSMLLCSLSEALCFPLGALTMPMPAVRTFALTSGLAVIDP 784
Db      526 CQNKRIPEDRTGECLKRTGASVALTSISNVTAFPAALIPALRAFSLQAQAVVVFNF 585
QY      785 LLQMSAFVALLSLDSKROEASRLDVCC-----VKQEL-----PPPG 822
Db      586 AMVLLIFPALLSMOLYRREDRLDIFCCFTSPCSRVIQVEPQAVTDTHDNTRYSPPPY 645
QY      823 QQEGGL----- 827
Db      646 SSHSFAHETQITMQSTVOLRTEYDPTHVYTTABRSEISVQPVTVTQDTLSQCSPEST 705
QY      828 -----LLGFFOKAYAPFLHMITRGGVLLFLFALFGVSL 861
Db      706 SSTRDLLSQSDSSHLCLPEPCKWTLSFPAEKHYAPFLPKPAKVVFVFLFGLGVSL 765
QY      862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEGAPVYFVTTGLNFSSEAGMNAICS 921
Db      766 YGTRVRDGLDLDIVPRETREYDFIAAQPKYFSF-----YNN----- 803

```


Db 973 FVEAIEKVRVVCNNYTSLSGSSYPNG-----YFP-----LFWEQYISLRHWLLLSISVVL 1022
QY 1121 VPTFAVSCILLGLDLASGLNLLSIYVILVDTVGFPMALWDISYNAVSLINLSAVGMSVE 1180
Db 1023 ACTFLVCAVFLNPNWTAGII-VNVLALMTVELFGWGLIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVFAVAVMTNLPGLILVLGLAKAQIQLF 1238
Db 1082 FTVHALAFLTAIGDKN---HRAMLALHMFAPVLDG-AVSTLLGLVLMLAGSEFDFIVRY 1137
QY 1239 FFRNLNLIITLLGLLHGLVFLPVLSVVG--PDVNPALALEQ---KRAEBAVAAMVASC 1293
Db 1138 FPAVALIITVLGVNLVLVLLSFFGCPCEVSPANGLNRLPTSPSPPPSVVRPAVPP 1197
QY 1294 NHPRSVS-TADNIYVNHSPGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISEBELROYEAQOGAG 1234

RESULT 18

US-08-656-055-10
; Sequence 10, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/656,055
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/540,406
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-656-055-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
QY 356 LVLSVTPVVALAAGLVFTLITDTPVELWSAPNSQARSEKAFHDQHPFPRTN-QVILTA 414
Db 88 LVVGLIFGAFAGVLKAANLETNVELWVEVGRVSRNELNYTRQKIGBEAMFNPQLMIQT 147
QY 415 PNRSSRYRYSLLGPKNFSGILDLDLLELLE--LQERLRHLQVMSPEAQRNLSQDICY 472

Db 148 PKBEG-----ANVLTTALLQHLDSALQASRVHVVMYN-----RQWKLHLCY 190
QY 473 AP-----LNPNTSLYDCCINSLSLQYFQNNRTLLLLTANQTLMGQ-----512
Db 191 KSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA--YLLGKPPPLRWNTNFDPL 248
QY 513 -----TSQVMDKHFLY-----CAN-----APL-----530
Db 249 BPLEELKKINYQVDSWEEMLNKAEBVGHGYMDRPNCLNADPCDPATAPKNKSTKPLDVALV 308
QY 531 -----TFKDG-----ALALSCMADYGAPVPFPPLAIGYKKG 562
Db 309 LNGGCGQLSRKYMHWQBELIVGTVKNATGKLVSALQTMFLQMTPKQMYHFHRYG---365
QY 563 DYSEABALIMTFLSNYPAGDPRLAQAKLWBEAFLEEMRAPQRRMAGMFQVTFTAERSLE 622
Db 366 DY-----VSHINWE-----DRAAILEAWQRTYVEVHQSVAPNSTQKVLPTT--TTLD 414
QY 623 DEINRTTAEPLIPATSYIVIFLYISLALGSYSWSRVNVDKATIGLGLGVAVVLGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMAYACLTM---LRWD---CSKSGAGVGLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSSLVILQVVPFLVSVGADNIPFVLEYORLPRRPGEPREPHVIGALGR 742
Db 470 GLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLLAHAFSETGQNKRIPFEDRTGCLXR 529
QY 743 VAPSMLLCSLSEACIFFLGALTMPNAVRTFALTSLAVILDLLQMSAPVALLSLDSKQ 802
Db 530 TGASVALTSISNVTAFMAALIPIPALRAFSLQAAVVVVFNFAMVLLIIPAILSMDLYR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPQGGGL-----827
Db 590 EDRRLDIFCCPTSPCVSRVIOVEPOAYTEPHSTRYSPPPPYTSHSFAHETHITMQSTVQ 649
QY 828 -----827
Db 650 LRTEYDPHTVYVTTAEPRSEISVQPVTVTQDNLSQSPSESTSTRDLSQFSDSLHCL 709
QY 828 -----LLGPFQKAYAPELLHWITRGVVLILLALFGVSLYSMSCHISVGLDOELALPK 879
Db 710 EPPCTKWTLSFPAEKHYAPFLLPKAKVVVILLFGLLGVSLYGTTRVRDGLDLTDIVR 769
QY 880 DSYLDYFLNRYEVEGAPVYFVTTILGYNFSRAGMAICSSAGCNSFTQKIQYAT- 938
Db 770 ETRYDFIAAQKYPF-----YNM-----YIVTKADYPMI 801
QY 939 -----EFPEQSYLAIPASS-----WVDDPIDWL-----TPSS---965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQMWLHYFRDWLQGLQDAPDSMETGRIMPNYKN 861
QY 966 -----CCRLYISGPNKDKPCPSTVNSLNCNCHSITMGSVRSPVSEQPHKYL-PWFLN 1017
Db 862 GSDDGVLAYKLLAVQTGRDK--PIDISQLT--KORLVADADGINFSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGGLAAY--STSVNLSTDGOVLASRFMAYHKPLKNSOD 1061
Db 916 DPVAYAAQANIRPHRPWVHDK---ADYMETRLRIPAAEPIEYAQPFYINGLRTSD 972
QY 1062 YTEALRAARELAANITA-DLRKVPCTDPAFEVFPYITNVFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEAIEKVRVVCNNYTSLSGSSYPNG-----YFP-----LFWEQYISLRHWLLLSISVVL 1022
QY 1121 VPTFAVSCILLGLDLASGLNLLSIYVILVDTVGFPMALWDISYNAVSLINLSAVGMSVE 1180
Db 1023 ACTFLVCAVFLNPNWTAGII-VNVLALMTVELFGWGLIGIKLSAVPVVILLASVIGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVFAVAVMTNLPGLILVLGLAKAQIQLF 1238
Db 1082 FTVHALAFLTAIGDKN---HRAMLALHMFAPVLDG-AVSTLLGLVLMLAGSEFDFIVRY 1137
QY 1239 FFRNLNLIITLLGLLHGLVFLPVLSVVG--PDVNPALALEQ---KRAEBAVAAMVASC 1293

Db 1138 FPAVALITVLGVLNGLVLLPVLLSFFGCPCEVSPANGNLRLPTSPPEPPSVVREAVPP 1197
Qy 1294 NHPSRVS-TADNIYNHSPFEGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELRQYEAQOGAG 1234

RESULT 19
US-08-954-668-10
; Sequence 10, Application US/08954668
; Patent No. 6172200
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,668
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.06
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-668-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.3%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

Qy 356 LVLSVIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHQHPFFRTN-QVILPA 414
Db 88 LVVGLLIFGAVGLKAALETNVEELMVEGVRSRELNYYTKQIGEEAMFNPLMIQT 147
Qy 415 PNRSSRYDILLGPNKFNFSGILDLLLELLE--LQERLHLQVWSPFQARNISLQDICY 472
Db 148 PKEEG-----ANVLTEALLQHLDSALQASRVHVVMYN-----RQWKLEHLCY 190
Qy 473 AP-----LNPDTSLVDCCINSLLQVFQNNRTLLLTANQTLMGQ----- 512
Db 191 KSGELITETGYMDQIETLYPLCLITPLDCLFWGAKLQSGTA--YLLGKPLPLRWTFDPL 248
Qy 513 -----TSQVDWKDHFLY-----CAN-----APL----- 530
Db 249 EFLEELKKINYQVDSWEMLNKAEGVGHYMDRFLCLNPADPCFATPNKNTKPLDVALV 308
Qy 531 -----TFKQGT-----ALALSCMADYGAPVFPFLAIGYKXG 562
Db 309 LNGGCGLSRKYMHQBELIVGGTVKNATGKLVSAHALQTMFLQMTPKQMYEHRGY--- 365

Qy 563 DYSEAEALIMTFGLNNYPAGDPRLPQAOKLWEEAFLEEMRAFORMMAGMFQVTFTAERSLE 622
Db 366 DY-----VSHINWNE-----DRAAAILEAQRTYVEVVHQSVAPNSQKVLPTT-ITLD 414
Qy 623 DEINRTAEDLPATSYIVIFLYISALGSYSWSRWVDSKATLGLGGVAVVLGAVMA 682
Db 415 DILKSPDSVIRVASGYLLMLAYACLTM--LRWD--CSKSGAVGLAGVLLVALSVAA 469
Qy 683 AMGFPSYLGIRSSILVILQVVPFLVLSVGADNIFVLEQYQRLPRRCPGEPEVHIGRALGR 742
Db 470 GLGLCSLIGISFNAATQTVLFFLAGVGVDVDFLLAHAFSETGQNKPIPEDRTRGECLKR 529
Qy 743 VAPSMLLCSLSEACFFELGALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSDSKRQ 802
Db 530 TGAVALTSISNVTAFFWAALIPALRAFSLQAAVVVFNFAVLLIFFAILSMDLVRR 589
Qy 803 BASLDVCC-----VKQEL-----PPQGEGL----- 827
Db 590 EDRLDIFCCFTPCVSRVQVEPQAYTEPHSNTRYSPPPPYTSHSFAHETHITWQSTVQ 649
Qy 828 ----- 827
Db 650 LRTEYDPHTHYTTAEPRSEISVQPVTVTDQNLSCQSPSTSTRDLLSQFSDSLHCL 709
Qy 828 -----LLGFFQKAYAPFLHWTITRGVLLLFALFGVLSYMSCHISVGLDDELALPK 879
Db 710 EPPCTKWTLSFAEKHYAPPLKPKAKVVVILLFLGLLSVLYGTTRVRDGLDITVPR 769
Qy 880 DSYLLDYFLNRYFEVGAQVYVFTTILGYNFSSEAGMNAICSSAGCNFSTQKIQYAT- 938
Db 770 ETREYDFAAQKYFSF-----YNN-----YIVTQKADYENI 801
Qy 939 -----EPPEQSILAIPASS-----WVDDFIDWL-----TPSS--- 965
Db 802 QHLLYDLHKFSNVKYVMLEENKQLPQWLHYFRDMLQGLQADPDSQWETGRIMNNYKN 861
Qy 966 -----CCRLYISGPNKDFCPTVNSLNCMLKNCMSITMGSVRPSVQFHKYL-PWFLN 1017
Db 862 GSDDGVLAYKLLVQTGSRDK--PIDISOLT--KQRLVDADGIINPSA--FYIYLTAMVSN 915
Qy 1018 D-----RNIKCPKGGLAAY--STSVNLTSDGVLASREWAYHKPLKNSOD 1061
Db 916 DPVAYAAQANIRPHRPEWHDK--ADYMPETRLRIPAAEPIEYQAQPFYINGLURDTS 972
Qy 1062 YTEALRAARELAANITA-DLRKVPCTDPAFVFPFYTTITNVFYEQYLTILPEGLFMLSCL 1120
Db 973 FVEAIEKRVVICNNYTSGLSSYPNG-----YFP-----LFEWQYISLRHWLLLSISVVL 1022
Qy 1121 VPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFPMALWDISYNAVSLINLVSAYGMSVE 1180
Db 1023 ACTFLVCVAVFLNPPWTAGII-VMVILAMTVLFGMMLGIGIKLSAVPVVILLIASVIGVE 1081
Qy 1181 FVSHITRSP--AISTKPTWLERAKETISMGSAPVAGVAMTNLPGLVILGLAKAQLIQIF 1238
Db 1082 FTVHVALAFLTAIGDKN--HRAMLALEHMFAPVLGD-AVSTLLGLVLMAGEPFIIVRY 1137
Qy 1239 PFRMLLLTLLGLHGLVFLPVILSYVG--PDVNPALALEQ---KRAEEAAVAVMVASCP 1293
Db 1138 FPAVALITVLGVLNGLVLLPVLLSFFGCPCEVSPANGNLRLPTSPPEPPSVVREAVPP 1197
Qy 1294 NHPSRVS-TADNIYNHSPFEGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELRQYEAQOGAG 1234

RESULT 20
US-08-918-658-10
; Sequence 10, Application US/08918658
; Patent No. 6429354
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L

TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,658
FILING DATE: 22-Aug-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/656,055
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:
LENGTH: 1434 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 10:

US-08-918-658-10

Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservativity 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVVPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFDQHPGPFRTN-QVILTA 414
DB 88 LVGLLIFGAPAVGLKAANLTVNELNVEGVGRVSRRLNTRQKIGBEAMFNPLMIQT 147
QY 415 PNRSSVRYDSLILGPKNFSGIIDLLELLE--LQERLRHLQVMSPEAQRNLSQDICY 472
DB 148 PKBEG-----ANVLTTEALLQHLDSALQASRVHVMYN---RQWGLEHLCY 190
QY 473 AP-----LNPDNTSLYDCINSLLQYFQNNRTLLILLTANQTLMOQ----- 512
DB 191 KSGELITETGYMDQITTEIYLPCLITPLDCFWEGAKLQSGTA--YLLGKPLPLRWTFDPL 248
QY 513 -----TSVDNKHDFLY-----CAN-----APL----- 530
DB 249 EFLBELKKINYQVDSWEEMLNKAEVGHGYNDRPCLNPADPCPATAPNKNSTKPLDVALV 308
QY 531 -----TFKDG-----ALALSCMADYGAPVFPFLAIGYKKG 562
DB 309 LNGGCGQLSRKYMHWQDEELIVGTVKNATGKLVSAHALQTMFLQMTFKQMYEHFRGY--- 365
QY 563 DYSEBALMTWTSNLNYPAGDPRLAOKLWEBAFLSEMRAPQRMAGMFOVTFATERSLE 622
DB 366 DY-----VSHINWNE---DRAAAILEAQRTYVEVHQSVAPNSTQKVLPEFTT--TTLD 414
QY 623 DEINRTTAEDLPFATSYVIFLYISLALGSYSYSSRVVVDKATGLGGAVALGAYMA 682
DB 415 DILKPSDVSIVRASGYLLMLAYCLTW---LRMD--CSKSGQAVGLAGVLLVALSVAA 469
QY 683 AMGFTSYLGIRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPRRPGEPREVHIGRALGR 742

DB 470 GLGLCSLIGISFNAATTQVLPLALGVGDVDFELLAHAFSETQCNKRIPFEDRTGCLCKR 529
QY 743 VAPSMILCSLSSEALCFPLGALTMPAVRTFALTSLGVLDFLLQMSAFVALLSLDSKQC 802
DB 530 TGAVALTISINVTAFMAALPIPALRAPLSQAQAVVVVFNFAWVLLIPAILLSMDLYRR 589
QY 803 EASRLDVCCC-----VKQBEL-----PPPGQCEGL----- 827
DB 590 EDRRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPYTSHTSHFAHETHITMSTVQ 649
QY 828 ----- 827
DB 650 LRTEYDPHTHYVYTTAEPRSEISVOPVTVTDNLSCQSPSTSTRDLSQFSSSLHCL 709
QY 828 -----LLGPFQKAYAPFLHMTITGVVLLLELALFGVSLYSMSCHISVGLDQELALPK 879
DB 710 BPCCTWTLSSFAEKHYAPFLKPKAKVVILLFLGLLGVSLYGTTRVRDGLDLTDIVPR 769
QY 880 DSYLLDYFLNRYFEVGPVYFVTTLTGYNFSSEAGMNAICSSAGCNFSFTQKIQYAT- 938
DB 770 ETRVEDFTAAQPKYFSF-----YNM-----YIVTQKADYPMI 801
QY 939 -----EPPEQSYLAIPASS-----WDDFDIWL-----TPSS--- 965
DB 802 QHLLYDLHKFSNVKYVMLEENKQLPQMWLHYFRDWLQGLQDAFSDSDWETGRIMPNNYKN 861
QY 966 -----CCRLVISGPNKDKCPCSTVNSLNCILKNCMSITWGSVRSPSVQPHKYL-PWFLN 1017
DB 862 GSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWVN 915
QY 1018 D-----RPNIKCPKGLAAY--STSVNLTSQGVGLASRFMAYHKPLKNSQD 1061
DB 916 DPVVAASQANIRPHRPENWHDK---ADYMPETLRIPAAEPIEYAFQFFYLLNGLRDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPGTDPAPFVFPYTTINVFYEQYLTILPGLFMLSLCL 1120
DB 973 FVEATEKVRVICNNVTSLGLSSYPNG-----YFP-----LPWEQYISLRHMLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLRSGLNLLSIVMLIVDTGVFMALWDISYNAVSLNLSAVGMSVYE 1180
DB 1023 ACTFLVCVAVFLNPNMTAGII-VMLALMTVELFGMGLIGIKLSAVPVVIVTASVIGIVE 1081
QY 1181 FYSHITRSF--AISTKPTWLERAKEATISMGSAGVAGVAMTNLPGLILVLGAKAQLIQIF 1238
DB 1082 FTVHVALAFLTAIGDKN---HRAMLALHMPAPVLDG-AVSTLLGVMLAGSEFDFIVRY 1137
QY 1239 PFRMLLITLLGLHLGVLPLVSYVG--PDVNPALALEQ---KRAEBAVAAMVVASCP 1293
DB 1138 PFAVLAAILTVLGLVGLVLLPVLSSFFGCPCEVSPANGLNRLTPSPBPSPSVVRFAVPP 1197
QY 1294 NHPSRVS-TADNIYVNHSPGSI-----KGAG 1319
DB 1198 GHTNNGSDSDSEYSSQTTVSGISEELRQYEAQOQAG 1234

RESULT 21

US-09-724-631-10
Sequence 10, Application US/09724631
Patent No. 6551782
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
JOHNSON, RONALD L
GOODRICH, LISA V
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/724,631
FILING DATE: 28-NO. 6551782-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/656,055
FILING DATE: 1996-05-31
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 1434 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-724-631-10
Query Match 9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity 22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;
QY 356 LVLSVTPVVALAGLVFTLTTPDVELMSAPNSQARSEKAFHQHGFPPFRN-QVILTA 414
DB 88 LVGLIFGAFAGVLAANLETVNELWEVGVRSRELYTRQKIGEMENPQLMIQT 147
QY 415 PNRSSRYSLIAGPKNFSGILDLDLLELLE--LQERLRHQLQVMSPEAQRNISILODICY 472
DB 148 PKEEG-----ANVTTEALLQHLDSALQASRVHVMYN-----RQWKLEHLCY 190
QY 473 AP-----LNPNTSYDCCINSLQYFQNNRTLLILLTANQTLMGQ----- 512
DB 191 KSGELITETGYMDQIIEYLIPCLIIITPLDCFWEGAKLQSGTA--YLLGKPLRLWTNFDPL 248
QY 513 -----TSQVDWKDHFLLY-----CAN-----APL----- 530
DB 249 EFLEELKKINYQVDSWEMLNKAEGHGYMDRCLNPADPCPATAPNKNSTKPLDVALV 308
QY 531 -----TFKDG-----ALALSCMADYGAPVFPFFLAIGYKKG 562
DB 309 LGGCCQLSRKYMHWQBELIVGTVKNATGKLVSAHALQTMFQMTPKQMYEHFRGY--- 365
QY 563 DYSEAEALIMTFSINYPAGDPLRAQAKLWEEAFLEEMAFQRRMAGMQVITFAERSLE 622
DB 366 DY-----VSHINWNE-----DRAAAILEAWQRTYVEVHQSVAPNSTQKVLPEPT--TLLD 414
QY 623 DEINRTTAEDLPITFATSYIVIFLYISALGYSYWSRVMVDSKATGLGGAVALVAVMA 682
DB 415 DILKPSDVSIVRASGYLLMLAYACTM---LRWD---CSKSGAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSSILVILQVVPFLVLSVGADNIIFVLEYQRLPRRGEPREVHIGRALGR 742
DB 470 GLGLCSLIGISFNAATTQVLPFLALGVGVDDVFLAHAFSETGQNKRIPEDEGTCECLR 529
QY 743 VASPMILCSISEALCFPLGALTWPARTALTSGLAVIDLFLQWSAFVALLSLDSKQ 802
DB 530 TGAVALTISINVTAFMAALIPALRAFSLQAAVVVFNFAVNLIFPAILLSMDLYR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPFGQGEGL----- 827

590 EDRLDIPCCFTSPCVSRVQIPEQAYTEPHSNTRYSPPPYTSHPAETHITMOSTVQ 649
828 ----- 827
650 LRTEYDPHTHYTTAEPREISVQPVTVTQDNLSCQSPSTSTRDLLSQFSSSLHCL 709
828 -----LIGFFQKAYAPFLLHWITRGVVLLFLALFGVSLYSCHMSHISVGLDQELAPK 879
710 EPCTKTWTLSSPAEKHYAPFLLKPAKVVILLFLGLGVSLYGTTTRVRDGLDLDIVPR 769
880 DSYLLDYFLNRYFEVGAAPVYFVTTLGYNFSSEAGNAICSSAGCNCNFSTQKIQAT- 938
770 ETREYDFIAAQFKYFSP-----YNN-----YIVTKADYPNI 801
939 -----EPQSQSLAIPASS-----WDDDFDML-----TPSS--- 965
802 QHLLYDLHKSFSNVKYMLEENKQLPQMWLHYFRDMLQGLQADAPDSWETGRINPNYKN 861
966 -----CCRLYISGPNKDFCPSVTNSLNCNMSITMGSVRPSVEQFHXYL-PWFLN 1017
862 GSDDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGIINPSA--FYILTAWVSN 915
1018 D-----RNINIKCPKGGLAAY--STSVNLTSGQVLAASFMAHYHKLKNSQD 1061
916 DPVAYAAQANIRPHRPEWVHDK---ADYMPETRLRIPAAEPTEYAQFPFYINGLRDTS 972
1062 YTEALRAARELANITA-DLRKVPGTDPAPFVPPYTTITNVFEQYLTILPEGLFMLSCL 1120
973 FVEAIEKVRVICNNYTSIGLSSYPNG-----YFP-----LFWEQYISLRHWLLLSISVVL 1022
1121 VPTPAVSCLLIGLDRSGLNLLLSIVMLVDTVGFMAIDISYNNAVLINLVSAVMSVE 1180
1023 ACTFLVCAVFLNFWTAGII--VMVLALMTVELFGMWGLIGIKLSAVPVPVILIASVGIVE 1081
1181 FVSHITRSP--AISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLQIF 1238
1082 FTVHVALAFLTAIGDKN---HRAMLALEHMPAPVLDG--AVSTLLGLVLMLAGSEDFIVRY 1137
1239 FFRNLNLTITLGLAHGLVFLPVLISYVG--PDVNPALALEQ---KRAEEAVALVAVASCP 1293
1138 FFAVALIUTVLGVNLGLVLLPVLISFPFGCPCEVSPANGNLRLPTPSPFPFVVRFAVPP 1197
1294 NHPRSVS--TADNIYVNHSPFESI-----KGAG 1319
1198 GHTNNGSDSSDSEYSSQTTVSGISEELRQYEAQOAG 1234
RESULT 22
US-08-954-701A-10
Sequence 10, Application US/08954701A
Patent No. 6610507
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley, Hoag & Eliot
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,701A
FILING DATE: 20-OCT-1997
CLASSIFICATION: 435

; TOPOLOGY: linear		
; MOLECULE TYPE: protein		
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:		
US-09-754-032-10		
Query Match		9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity		22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
Qy	356	LVLVSVVVAALAGLVFTLTDPVELMSAPNSQARSEKAFHQHFGFPFRFN-QVILTA 414
Db	88	LVVGLLIIFGAFVGLKAANLETNVEELNVEGVRSRELNTRYTRQKIGEEAMFNPLMIQT 147
Qy	415	PNRSSRYVDSLLGPKNFSGILDLLLELLE--LQERLHLQVMSPEAQNISLQDICY 472
Db	148	PKREG-----ANVLTTEALLQHLDSALQASRVHYMYN-----RQWKLEHLCY 190
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQ----- 512
Db	191	KSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA--YLLGKPLRLWTFDPL 248
Qy	513	-----TSQVDMKDHFLY-----CAN-----APL----- 530
Db	249	EFLEELKKINYQVDSWEEMLNKAENVGHGYMDRPLNADPCDPATAPNKNSTKPLDVALV 308
Qy	531	-----TFKDGTT-----ALALSCMADYAGPVPFPFLAIGYKKG 562
Db	309	LNGGCCGLSRKYMHWQBELIVGTVKNAATGKLVSAHALQTMFQMTPKQWYEHFRGY--- 365
Qy	563	DYSEAEALIMTFTSLNNYPAGDPRPRLAQAKLWEERAFLEEMRAFORRMAGMFOVTFTAERSLE 622
RESULT 24		
US-08-916-140-10		
Sequence 10, Application US/08916140		
Patent No. 6946257		
GENERAL INFORMATION:		
APPLICANT: Scott Mathew P.		
APPLICANT: Goodrich, Lisa V.		
APPLICANT: Johnson, Ronald L.		
APPLICANT: Epstein, Ervin Jr.		
TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO		
FILE REFERENCE: CIBT-P04-203		
CURRENT APPLICATION NUMBER: US/08/916.140		
CURRENT FILING DATE: 1997-08-21		
PRIOR APPLICATION NUMBER: US 08/656,055		
PRIOR FILING DATE: 1996-05-31		
PRIOR APPLICATION NUMBER: US 08/540,406		
PRIOR FILING DATE: 1995-10-06		
PRIOR APPLICATION NUMBER: US 08/319,745		
PRIOR FILING DATE: 1994-10-07		
NUMBER OF SEQ ID NOS: 64		
SOFTWARE: FastSeq for Windows Version 4.0		
SEQ ID NO 10		
LENGTH: 1434		
TYPE: PR		
ORGANISM: Mouse		
US-08-916-140-10		
Query Match		9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity		22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
Qy	356	LVLVSVVVAALAGLVFTLTDPVELMSAPNSQARSEKAFHQHFGFPFRFN-QVILTA 414
Db	88	LVVGLLIIFGAFVGLKAANLETNVEELNVEGVRSRELNTRYTRQKIGEEAMFNPLMIQT 147
Qy	415	PNRSSRYVDSLLGPKNFSGILDLLLELLE--LQERLHLQVMSPEAQNISLQDICY 472
Db	148	PKREG-----ANVLTTEALLQHLDSALQASRVHYMYN-----RQWKLEHLCY 190
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQ----- 512
Db	191	KSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA--YLLGKPLRLWTFDPL 248
Qy	513	-----TSQVDMKDHFLY-----CAN-----APL----- 530
Db	249	EFLEELKKINYQVDSWEEMLNKAENVGHGYMDRPLNADPCDPATAPNKNSTKPLDVALV 308
Qy	531	-----TFKDGTT-----ALALSCMADYAGPVPFPFLAIGYKKG 562
Db	309	LNGGCCGLSRKYMHWQBELIVGTVKNAATGKLVSAHALQTMFQMTPKQWYEHFRGY--- 365
US-10-736-769-4		
Query Match		9.4%; Score 646.5; DB 2; Length 1434;
Best Local Similarity		22.9%; Pred. No. 3.3e-51;
Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;		
Qy	356	LVLVSVVVAALAGLVFTLTDPVELMSAPNSQARSEKAFHQHFGFPFRFN-QVILTA 414
Db	88	LVVGLLIIFGAFVGLKAANLETNVEELNVEGVRSRELNTRYTRQKIGEEAMFNPLMIQT 147
Qy	415	PNRSSRYVDSLLGPKNFSGILDLLLELLE--LQERLHLQVMSPEAQNISLQDICY 472
Db	148	PKREG-----ANVLTTEALLQHLDSALQASRVHYMYN-----RQWKLEHLCY 190
Qy	473	AP-----LNPDNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQ----- 512
Db	191	KSGELITETGYMDQIIEYLYPCLIIITPLDCFWEGAKLQSGTA--YLLGKPLRLWTFDPL 248
Qy	513	-----TSQVDMKDHFLY-----CAN-----APL----- 530
Db	249	EFLEELKKINYQVDSWEEMLNKAENVGHGYMDRPLNADPCDPATAPNKNSTKPLDVALV 308
Qy	531	-----TFKDGTT-----ALALSCMADYAGPVPFPFLAIGYKKG 562
Db	309	LNGGCCGLSRKYMHWQBELIVGTVKNAATGKLVSAHALQTMFQMTPKQWYEHFRGY--- 365
Qy	563	DYSEAEALIMTFTSLNNYPAGDPRPRLAQAKLWEERAFLEEMRAFORRMAGMFOVTFTAERSLE 622
Db	366	DY-----VSHINWNE-----DRAAILEAWQRTYVEVHQSVAPNSTOKVLPFTT-TILD 414
Qy	623	DEINRTABDLPFATSYIVIFLYISLALGSYSRSMVMSVDSKATLGLGSVAIVLGAAMA 682
Db	415	DILKSFSDSVIRVASGYLLMAYACTM--LRWD--CSKQGAAGLAVLLVALSVAA 469
Qy	683	AMGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLEYQRIPLRPPGEPREHIGALGR 742
Db	470	GLGLCSLIGSFNAATQVLPFLALGVGDVDFLLAHAFSETGQNKRIPEFDRTEGLKR 529
Qy	743	VAPSMCLCSIEAICFELGALTPMPAVRTFALTGLAVILDFFLOMSAPVALLSDSKRQ 802
Db	530	TGASVALTSISNWTAFMAALIPALRAFSLQAAVVVFNFAWVLLIFPAILSMDLR 589
Qy	803	EASRLDVCCC-----VKPQEL-----PPQGQGL----- 827
Db	590	EDRLDIFCCFTSPCVSRVQVEFQVATBPHSNTRYSPPPVTVSHSFAHETHITMQSTVQ 649
Qy	828	----- 827
Db	650	LRTEYDPHTHYVTTAEPRISEISQPVTVTQDNLSCQSPSTSTRDLSQFSDSLHCL 709
Qy	828	-----LIGFFQKAYAPFLHWITRGVLLFLALFGVSLYSMSCHISVGLDQELALPK 879
Db	710	EPCTKTWLSFAEKHYAPFLPKAKVVVILLFLGLGLVSLGTTTRVRDGLDITDIVER 769
Qy	880	DSYLLDYFLNRYFEVGAPVYFVTTLGYNFSSAGMNAICSSAGCNFSTFKIQIYAT- 938
Db	770	ETREYDFIAQPKYFSF-----YNM-----YIVTQKADYPNI 801
Qy	939	-----BPPQSYLAIPASS-----WDDFDIML-----TPSS--- 965
Db	802	QHLLYDLHKFSNVKYMLENKQLPQMWLHYFRDMLQGLQADAFDSWETGRIMPNNYKN 861
Qy	966	-----CCRLYISGPNKDFCSTVNSLNCNCKMSITMGSVRPSVEQPHKIL-PWFLN 1017
Db	862	GSDGVLAYKLLVGTGRDK--PIDISQLT--KQRLVDAGIINPSA--FYIYLTAWSN 915
Qy	1018	D-----RENIKCPKGLAAY--STSVNLTSDGQVLASRFMAHYHKLKNSQD 1061
Db	916	DPVAYAAQANIRPHREWHDK--ADYMPETRLRIPAAEPIEYAFPPFLNGLRDTSD 972

Db 366 DY-----VSHINWE-----DRAAILEAWQRTVEVHOSVAPNSTQKVLPETT-TTLD 414
QY 623 DEINRTTAEDLPFATSYIYIFLYISLALGSYSWSRVVMDSKATLGLGGAVALGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMLAYACTM--LRWD--CSKSQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIPFVLEYQRLPRRGPGEPRVHIGALGR 742
Db 470 GLGLCSLIGISFNAATTVLPFLALGVGVDDVFLAHAFSETQKRIIPFEDRTGCLXR 529
QY 743 VAPSMILCSLSEALCFGLGALTMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLSKQ 802
Db 530 TGSVALTSISNTAFMAALIPALRAFSLQAQAVVVFVFAVLLIPAILSMOLYXR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPPGQEGEGL-----827
Db 590 EDRRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPYTHSHFAHETHITMQSTVQ 649
QY 828 -----827
Db 650 LREYDPHTHVYVYTTAEPRSEISQPVTVTQDNLSQSPSTSTRDLSQFSDSLHCL 709
QY 828 -----LLGPFQKAYAPLHAWITRGVLLFLFALFGVSLYSMCHISVGLDOELALPK 879
Db 710 EPPCTKWTLSFAKHYPAPLKPAAKVVVLLFLGLLGVSLYGTTRVRDGLDLDIVPR 769
QY 880 DSYLLDYFLNRYFVPGVAVPVVTVTLGYNFSSEAGMNAICSSAGCNFSFTQKIYAT- 938
Db 770 ETRYDFIAAQPKYFSP-----YNM-----YIVTQKADYPMI 801
QY 939 -----EPFEQSYLAIPASS-----WDDPIDWL-----TPSS-----965
Db 802 QHLLYDLHKSFSNVKVMLEENKQLPQWMLHYPRDMLQGLQDAFSDWETGRIMPNNYKN 861
QY 966 -----CCRLYISGNKDKFCEPSTVNSLNCNCHSIYMGSVRPSVEQPHKYL-PWFLN 1017
Db 862 GSDGVLYAYKLLVQTSRDK--PIDISQLT--KQRLVDADGGINPSA--FYIYLTAWVSN 915
QY 1018 D-----RPNIKCPKGGAAAY--STSVNLTSDBGVLAASRFWAYHKPLKNOD 1061
Db 916 DPVAVASQANTIRPHPEWHDK--ADYMPETRLRIPAEPIEYIAQFFPYLNGRLDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPGTDPAPFEPYITTVNVEQYLTILPEGLFMLSCL 1120
Db 973 FVEATEKRVAVICNNVTSLSGLSVNG-----YFP-----LFWEQYISLRHMLLSISVL 1022
QY 1121 VPTFAVSCLLGLDLBSGLNLLSIVMILVDTGFMALWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCAVFLNPMWTAGII-VNVLALMTVELFGMGLIGIKLSAVPVVILLIASVGIVE 1081
QY 1181 FVSHITRSP--AISTKPTWLERAKEATISGSAVAGVAMTNLPGLVILGLAKAQLIQIF 1238
Db 1082 FTVHVVALPFLTAIGDN-----HRAMLALHMPAPVLDG-AVSTLGLVLMLAGSEFPFIVRY 1137
QY 1239 PFRANLLITLLGLHLGLFPLVLSYVG--PDVNPALAEQ---KRAEEAVALVAVASCP 1293
Db 1138 PFAVLAAILTVGLNGLVLLPVLSPFGPCPEVSPANGNLRLPTSPPEPPSVVRFAVPP 1197
QY 1294 NHPRSVS-TADNIYVNHSPFESI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSSQTTVSGISBELROYEAQOAG 1234

RESULT 25

PCT-US95-13233-10

; Sequence 10, Application PC/TUS9513233
; GENERAL INFORMATION:
; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13233
; FILING DATE: 06-OCT-1990
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13233-10

Query Match 9.4%; Score 646.5; DB 4; Length 1434;

Best Local Similarity 22.9%; Pred. No. 3.3e-51;

Matches 283; Conservative 171; Mismatches 420; Indels 363; Gaps 44;

QY 356 LVLSVIPVVALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQDHFQPPFRN-QVILFA 414
Db 88 LVVGLLIIFGAFVAGKAAANLETNVBEVGVGRVSRRLNTRYQKIGBAMFNQMIQT 147
QY 415 PNRSSRYVDSLLGLPKNFSGILDLALLLELLR--LQERLRHLQVMSPEAQNISLODICY 472
Db 148 PKBEG-----ANVTTEALQHLDSALQASRVHVVMYN-----RQWLEHLCY 190
QY 473 AP-----INPNTSLYDCCINSLLQYFQNNRTLLLLTANQTLMGQ-----512
Db 191 KSGELITETGYMDQIIIEVLYPCLIIITPLDCFEGAKLQSGTA--YLLGKPLRLWTNFDPL 248
QY 513 -----TSQVDWKDHFY-----CAN-----APL-----530
Db 249 EFLEBLKKINQVDSWBEMLNKAARYGHGYMDRPLNPDPCPATAPNKNSTKPLDVALV 308
QY 531 -----TFKQGT-----ALALSCMADYAGAPVPPFLAIGSYKQK 562
Db 309 LNGGCGQLSRKTMHQBELIVGGTVKNATGKLVSAHALQTMPLQMTPKQMYEHFRGY----365
QY 563 DYSEALALIMTFSLNNYPAGDPRLLAQAKLWBEAFLEEMRAFORRMAGMFQVFTFAERSLIE 622
Db 366 DY-----VSHINWE-----DRAAILEAWQRTVEVHOSVAPNSTQKVLPETT-TTLD 414
QY 623 DEINRTTAEDLPFATSYIYIFLYISLALGSYSWSRVVMDSKATLGLGGAVALGAVMA 682
Db 415 DILKSFSDSVIRVASGYLLMLAYACTM--LRWD--CSKSQAGVLAGVLLVALSVAA 469
QY 683 AMGFPSYLGIRSLVILQVVPFLVSVGADNIPFVLEYQRLPRRGPGEPRVHIGALGR 742
Db 470 GLGLCSLIGISFNAATTVLPFLALGVGVDDVFLAHAFSETQKRIIPFEDRTGCLXR 529
QY 743 VAPSMILCSLSEALCFGLGALTMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLSKQ 802
Db 530 TGSVALTSISNTAFMAALIPALRAFSLQAQAVVVFVFAVLLIPAILSMOLYXR 589
QY 803 EASRLDVCCC-----VKPQEL-----PPPGQEGEGL-----827
Db 590 EDRRLDIFCCFTSPCVSRVIOVEPQAYTEPHSNTRYSPPPYTHSHFAHETHITMQSTVQ 649

```
QY 828 ----- 827
Db 650 LRTEYDHPHTYTTAPRSEISQPVVTQDNLSQSPSTSTRDLSQFSSLSHCL 709
QY 828 -----LLGFFOKAVAPFLHWTITGVVLLFLALFGVSLYGMCHSISVGLDQELALPK 879
Db 710 EPPCTKWTLSFAEKHAPFLKPKAKVVVILLFGLGVSlyTTRVROGLDLDIVPR 769
QY 880 DSYLLDYFLNRYFEVGPVYVTTGLYNPSSBAGMNAICSSAGCNFSFTKIQIYAT- 938
Db 770 ETREVDPTAAQKYFSE-----YXN-----YIVTKADYENI 801
QY 939 -----EPQSYLAIPASS-----WVDDFTDWL-----TPSS--- 965
Db 802 QHLLYDLHKFSNRYKYNMLEENKQLPOMWLHYFRDWLQGLQADPDSOWETGRIMPNNYKN 861
QY 966 -----CCRLYISGPNKDFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYL-PWPLN 1017
Db 862 GSDGVLAyklVQTGRDK--PIDISQLT--KQRLVDADGIINPSA--FYIYLTAWSN 915
QY 1018 D-----RPNIKCPKGGLAAY--STSVNLTSDQVLAASRPMAHYKPLKNSQD 1061
Db 916 DPVAAASQANIRPHRPWVHDK--ADYMPETRLRIPAAEPYEAQPPFYLNGLRDTS 972
QY 1062 YTEALRAARELANITA-DLRKVPCTDPAPFVPPYITNVYEQYLTILPGLFMLSCL 1120
Db 973 FVEAIEKVRVICNNYTSIGLSSYPNG-----YFP-----LFWEQYISLRHWLLSISVVL 1022
QY 1121 VPTFAVSCLLGLDLSGLLNLSTVMILVDTPGPMALWDISYNAVSLINLVSAGMSVE 1180
Db 1023 ACTFLVCVFLNPTAGII-VWULAMTWELFGMGLIGIKLSNVPVILLIASVGIGVE 1081
QY 1181 FVSHITRSF--AISTKPTWLERAKETISMGSAVEPAGVAMTNLPGLVGLAKAQLIOIF 1238
Db 1082 FTVHVALAFLTAIGDKN--HRAMLAEHMAPVLGD-AVSTLLGLVLMAGSEDFIVRY 1137
QY 1239 PPRMLLTLLGLLHGLVFLPILSYVG--PDVNPALAEQ---KRAEVAANVMVASCP 1293
Db 1138 FFAVLAAILTVLGLNGLVLLPVLFFGFCPEVSPANGLNRLPTSPSPFPVVRFAVPP 1197
QY 1294 NHPRSVS-TADNIYVNHSPFSGSI-----KGAG 1319
Db 1198 GHTNNGSDSDSEYSQTTVSISRELQRYEAQQAG 1234

RESULT 26
US-09-248-796A-16161
; Sequence 16161, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16161
; LENGTH: 642
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16161

Query Match 9.3%; Score 639.5; DB 2; Length 642;
Best Local Similarity 26.4%; Pred. No. 3.7e-51;
Matches 188; Conservative 123; Mismatches 291; Indels 111; Gaps 21;
QY 6 LRGWLLWALLRLAQSEPYTTIHQPYGCAPYDCGKNPGLSGSLMTLSNVSCLSNTPARK 65
```

```
Db 12 MRAICLLLLTINLAIAS-VLSLHKPGYCNTRYGNCCKKSVFGKPL-----PCAEEFPAVK 64
QY 66 ITGDHILLLOKICPRIPTYGNTQACCSAKOLVSLASLSITKALLTFCPCASDNFVNLHC 125
Db 65 ASQESREKLKISCGKDF-----DYICCSPEQIDILELNLKRVDPILISSCPACRKNFYDFPC 120
QY 126 HNTCSNOSLFINVTRVAQLGAGQAPAVAYEAFYQHSFAEQSYDSCSRVRVPAATLAV 185
Db 121 QFSCSPNESQFVEIITKETARDTGKEIVTEINQYVEPGMANQFPDSCKNKFLATN--- 176
QY 186 GTMCGVYGSAALCNAQRWLNFGQDTGN--GLAPLDITPHLLEPGQAVGSGLOPLNEGVARC 243
Db 177 GYAMDLLGGGAKNVQLQKPLGDEKPLLLGGSPYQINEVYKLP--ETDSGLVLRNEPLRDC 234
QY 244 NESQGDVATCSCODCAASCPATAPQALDSTFVLGOMPSGLVLIILILCSVFAVVITLLV 303
Db 235 NDKE-----YKCACTDCBESCPKLPFAKDLTKCTGVGLPCFSFSIII---IWSCHVILG 287
QY 304 GERVAPAR-----DKSKWDPKKGTSLSDKLSFSFTHLLGQFQGW---- 344
Db 288 GYHVLAKLKERRRSIAEDSEDDSTMINPLFYAGLGKRAKQFLSEIGLKIOWFANI 347
QY 345 GTWVASWPLTILVLSVIPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHQHGFPP 404
Db 348 GYFCSKEFGISIGTSLAVVLLSLGLFLKQLETPVKLVSPNDPAYKNQYFESFGEW 407
QY 405 FRTNQVLTAPNRSSRYDSLLGPKNFSGLDLDLLELLELQERLRLHQVMSPEAQRN 464
Db 408 FRIEQVTVSS-----KDDGVLNWDIVKVMWFDKESQLETL-----NEN 445
QY 465 ISLDQICVAPLNPDNTSLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDWKDHFY 524
Db 446 VRLSDICFKPL--DET---CALQSFQYFGD-----ISGLT-ETNWSKLQS 487
QY 525 CANAPLTFKGTALALSCMADYAGVPFPFPAIGYKDKYSEALIMTFSLNNYPAGDP 584
Db 488 CVDSF-----VNCLPTFQQLKPNIL---FDSNDISQAKAFTTVLVNSDTQNEH 534
QY 585 RLAQAKLWEAFLEEMRAFORRMAGMQVTTAERSLEDEINRTTAEPLPIFATSYIVIF 644
Db 535 YTSNTISYEHLPQKWAADLQTEYNL-NTAYSTSILKEELNQSNTDIKTIALSYLVWF 593
QY 645 LYISLALGS-----YSSMSRVVDSKATILGLGGVAVVLGAVMAAMGFFSYL 690
Db 594 IYASLALGGLKPSANLYS-----LVKTFEFTLGFSSIIILLSVTASAGSFLLL 641

RESULT 27
US-08-857-636-60
; Sequence 60, Application US/08857636
; Patent No. 6552181
; GENERAL INFORMATION:
; APPLICANT: Dean, Michael Carlton
; APPLICANT: Hahn, Heidi Eve
; APPLICANT: Wicking, Carol
; APPLICANT: Christiansen, Jeffrey G.
; APPLICANT: Zaphiropoulos, Peter G.
; APPLICANT: Gailani, Mae R.
; APPLICANT: Shanley, Susan Mary
; APPLICANT: Chidambaram, Abirami
; APPLICANT: Vorechovsky, Igor
; APPLICANT: Holmberg-Lindstrom, Erika
; APPLICANT: Uden, Anne Birgitte
; APPLICANT: Gillies, Susan Alana
; APPLICANT: Negus, Kylie
; APPLICANT: Smyth, Ian Mcleod
; APPLICANT: Pressman, Carol Leah
; APPLICANT: Leffell, David J.
; APPLICANT: Gerrard, Bernard
; APPLICANT: Goldstein, Alisa Miriam
; APPLICANT: Wainwright, Brandon
; APPLICANT: Toftgard, Rune Carl-Magnus
```

APPLICANT: Chenevix-Trench, Georgia
APPLICANT: Bale, Allen E.
TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene
NUMBER OF SEQUENCES: 83
CORRESPONDENCE ADDRESS:
ADDRESSES: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/857,636
FILING DATE: 16-MAY-1997
CLASSIFICATION: 514
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/017,906
FILING DATE: 17-MAY-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: AU P00011
FILING DATE: 21-MAY-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: AU P00363
FILING DATE: 07-JUN-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 60/019,765
FILING DATE: 14-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Hunter, Tom
REGISTRATION NUMBER: 38,498
REFERENCE/DOCKET NUMBER: 015280-278200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..1296
OTHER INFORMATION: /note= "amino acids encoded by human
nevoid basal cell carcinoma syndrome
(NBCCS) (PATCHED (PTC)) CDNA"
US-08-857-636-60

Query Match 9.1%; Score 628; DB 2; Length 1296;
Best Local Similarity 23.6%; Pred. No. 1.6e-49;
Matches 254; Conservative 165; Mismatches 384; Indels 272; Gaps 39;

QY 426 LIG--PKNFGSLDLDLLELLELLELLELLOVMSPEAQRNLS-----LQDICYAPLNP 477
DB 97 LLGKPLKRWTFDPLELELKKI-----NQYDSWEEMLNKAABVGHGMDRPLCLNPADP 151
QY 478 DNTSLYDCINSLLOFYQNNRFLLLLTANQTLMGQTSQ-VQWKDHFYCANAPLTFKDG 536
DB 152 D-----CPATAPKNSTKPLDMALVNGCGHLSRKMYHWEELIVGG-----TVKST 200
QY 537 -----ALALSCWADYGAVPFPPLTATGGYKDKYSEARALIMTFSLNNYPAGDPRLAQAKL 591
DB 201 GKLSAHALQTMFLMTPKQMY---EHPKGYEY-----VSHINWNE-----DKAAAIL 247
QY 592 WEEAFLERAPORRMAGHFQVTFPAERSLEDEINRTTAEDLPFATSYIVIFLYISAL 651
DB 248 WORTYVEVHOSVAQNSTQKVLSTT-TLDDILKSPSDSVIRVAGYLMLLAYACITM 306

QY 652 GSYSSWSRVMVDSKATLGLGAVVAVLGAVMAMGFFSYLGRSSLVILQVVPFLVLSVGA 711
DB 307 ---LRWD--CSKSGAGVAGLAVLLVALSVAAAGLGLCSLIGISFNAATQVLPLPGLGVG 361
QY 712 DNIPIFVLEYQRLPRPGEPREPHVHIGRAIGRVAPMSMLLCSLSEATCFPGALGATPMPAVRT 771
DB 362 DDVFLAHAFSETGQNKRIPEPDRTGECILKRTGASVALTSISNVTAFPMALPIPALRA 421
QY 772 PALTSLGAVILDFLLQMSAFVALLSLDSKROEASRLDVCC-----VKPQEL-- 818
DB 422 FSLQAAVVVVFNFAMVLLIFPAILSMDLYRRDRDLDFCCFTSPCVSRVIOVEPOAYTD 481
QY 819 -----PPPGQGEGL----- 827
DB 482 THDNTRYSPPPPYSSHSPAHETQITMQSTVQLRTBYDHPHVVYTTAEPRSSISVQPVTV 541
QY 828 -----LIGFPKAYAPFLHLHMITRGV 848
DB 542 TQDTLSQSPSTSTRDLSQFSDSLHCLPEPPCTKWTLSFAEKHVAFLPKPKAKVV 601
QY 849 VLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDFLNLNRYFEVGAPVYFVTILGY 908
DB 602 VIFFLGLGLVSLYGTTRVGLDLDIVPRETRYDFIAAQKFYSP-----Y 650
QY 909 NFSSEAGMNAICSSAGCNFSPFQKIYAT-----EPPEQSILAIPASS-----W 953
DB 651 NM-----YIVTQADYPIQHLLYDLHRSFSNVKYVYMLBENKQLPKWM 693
QY 954 VDDFIDWL-----TPSS-----CCRLIYISGPNKDKPCPSTVNSL 987
DB 694 LHYPRDWLQGLQDAFDSWETGKIMPNNYKNGSDGVILAYKLLVQTGSRDK--PIDISOL 751
QY 988 NCLKNCSITMGSVRPSVEQHKYL-PWFLND-----RPNIKCPKGLAAY 1032
DB 752 T--KORLVADAGIINPSA--FYILTAWVSNDPVAYAASQANIRPHRPEWVHDK---ADY 804
QY 1033 --STSVNLTSDGQVLARFMAYHKPLKNSQDYTEALRAARELAANITA-DLRKVPQTDPA 1089
DB 805 MPETRLRIPAAEPIEYAGFPFVNLGLRDTSDVEAIEKVRTICSNYTSIGLSSYPNG--- 861
QY 1090 FEVPYTTNVFYEQYLITLPEGLPMLSLCLVPTPAVSCLLGLDLRGLMLLMLVMTL 1149
DB 862 ---YFP---LFWEQYIGLRHMLLIFISWLACTFLVCAVFLNPNWTAGII--VMVALMT 913
QY 1150 VDTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLERAKEATIS 1207
DB 914 VELFGMGLIGIKUSAVPVVILIASVGIVGVFTVHVALAFLTAISDKN---RRVLALEH 970
QY 1208 MGSVAFGAVMTNLPGLVLGLAKAQLTIQIFPRLNMLLITLIGLHGLVFLPVLISYVG- 1266
DB 971 MPAPVLDG-AVSTLLGVLMAGSEPDFIVRYFFAVLAILTILGVNLGLVLLPVLISFPGP 1029
QY 1267 -PDVNPALALEQ---KRAEAAVAVMVASCPNHPERSVSTADNIYVNHSPFSGIKG 1317
DB 1030 YPEVSPANGLNRLPTSPSPPPSVVRFAFMPGCHTH--SGSDSSDSEYSSQTTVSG 1082

RESULT 28
US-09-293-505-2
; Sequence 2, Application US/09293505
; Patent No. 6348575
; GENERAL INFORMATION:
; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 2

```
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-293-505-2

Query Match      8.5%; Score 584; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 2.1e-45;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;

Qy 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLLGQFFQG----WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLLFSLGCGIQRHCGKVLFLG 62
Qy 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHFQ--PFFRTNQVILTAPNR 417
Db 63 LLAFGALALGLRMAIETNLEQLWVEGSRVSEQLHYTKELGEEAAYTSQMLIQTAQOE 122
Qy 418 SSRYDSSLGLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQIDICY---AP 474
Db 123 GEN-----ILTPALG-----LHLQALTASKVQVSLYCKSWDLNKKIYKSGVP 166
Qy 475 LMPDN-----TSLYDCCINSLLQVFNRRITLLLTANOTLMGQTSQVDWKD----- 520
Db 167 LIENGMIEMIEKLPFCVILTLPDLCFWGAK---LQGSAYLPGRPDIQWTNLDPEQLLE 223
Qy 521 -----HFLYCANAPLTFKDGKTALALSCMADYGA 548
Db 224 ELGPFASLEGFRELDKAQVGYVRPCLHPDDLHCPSPAPNHHSRQAPNVVAHELSGGC 283
Qy 549 PVPEP-----LAIGYKGDYSE---AEALIMTF-----SLNNYPAGDPRLAQAKL 591
Db 284 HGFSHKFMWQBELLLGGWARPQOELLRAEALQSTFLLMSPRQLYEHFRGDYQTHDIGW 343
Qy 592 WEEAFLEENRAPQRRMAGMFQVTTAERSLE-DEINRTTAEDLPITATS-----Y 640
Db 344 SEEQASTVLQAWRRPVLQAEALPENASQIHFSSITLDDI-LHAFSEVSAARVWVG 402
Qy 641 IVIFLYISALGYSWSRWVDSKATGLGAVVILGNVMAAGFFSVGLIRSSILVILQ 700
Db 403 LLMLAYACVTM---LRWD---CAQSGSVGLAGVLLVALAVASGLGCLALLGITFNAATQ 457
Qy 701 VVPELVLSGADNIFVLEYQRLPRRCPGEPEVHIGRALGRVAPSMILCSLSEACIFFL 760
Db 458 VLPFLALGVDVDFL--LAHAFTALPGTPOERMGECLOQTGTSVVLTSINNMAAFILM 515
Qy 761 GALTMPAVRTFALTSLGLAVILDFLLQMSAFVALLSLSKQESARLNVCCC----- 812
Db 516 AALVPIPALRAFSLQAAIVVGCTFVAVMLVFPAILSLDRRRHCORLDVLCFCFSSPCSAQ 575
Qy 813 ---VKQEL-----PPPGQEGEL-- 827
Db 576 VIQILPQELGDTGTVPGVIAHLTATVQAFTHCEASSQHVITLPPQAHVLPFSPDPLGSEL 635
Qy 828 -----LLG-----PFQKAYAPFLHMITRQVLLPLFLA 855
Db 636 FSPGSTRDLGQEEETRQKACKSLPCARWNLHAEPYQFAPLLIQSHAKAIVLVLFCA 695
Qy 856 LFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVTVFTTLGNFS-SEA 914
Db 696 LLGLSLYGATLVQDGLATDVVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDVAHSQR 754
Qy 915 GMAIACSSACGNFSTQKIQYATEPPEQSYLEIPAS-----SWDDDFIDWLTPSSCC--- 967
Db 755 AL-----FDLHQRFSSL-----KAVLPPPTAQPRTHLYRNWLOGIQIAFPQ 798
Qy 968 -----RLYISGPNKOKFCPSTVNSLNCNKMCSITWGSVRPSVEQFHKYLPWFLND 1018
Db 799 DWASGRITRHSYNG-SED-----GALAYKLLIQTGAQEPID-FSOLITTRKLV 846
Qy 1019 RPNKCPKGLAAYSTSVNITSDGVLAASFMAHYKP----- 1055
Db 847 REGLIIPPE--LFYMGTLVWVSSDPLGLAASQANFPYPPPEWLHDKYDTTGENLRIPPAQ 904
```

```
1056 -----LKNQDYTEALRAARELAANI-TADLRKVPGTDPAPFVFPYTTNVFY 1102
905 LEPAQPPFLRLGLOKQTADFVEATEGARAAACABAGQAGVHAYPSGSP-----LFW 954
1103 EOYLITLPEGLFMLSIC--LVPTFAVSCLLGLDLRSGLNLLSIVMILYDVTGVPFMAWD 1160
955 EYVLGL--RRCFLLAVCILLVCTFLVCALLLLNLPWTAGLI-VLVLANMTVELFGINGFLG 1011
1161 ISYNAVSLINLVASGMSVEFVSHITRSFPAISKTPTWLERAKETISMGSVAVFAGVAMTN 1220
1012 IKUSAIPDVITLVASVGVVEFTVHVALGF-LTTQGSNRLRAHALEHTFAPVTOG-AIST 1069
1221 LPLGILVLGAKAQLIQIFPFRLLNLLIITLLGLLHGLVLPVLVSYGPDVNPALALEQRA 1280
1070 LLGLLMLAGSHDFIVRYFFAALTIVLTLGLLHGLVLLPVLLSILGPP--PEVIQMYKES 1127
1281 EEAVAAV-----MVASCPNHPSPSVSTADNIYN 1308
1128 PEILSPPAQGGGLRWGASSSLPQSFARVTTTMTVAIH 1165

RESULT 29
US-09-060-939A-2
; Sequence 2, Application US/09060939A
; Patent No. 6709838
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage, David A. Carpenter
; TITLE OF INVENTION: Patched-2
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060.939A
; FILING DATE: 15-Apr-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1203 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-060-939A-2

Query Match      8.5%; Score 584; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 2.1e-45;
Matches 284; Conservative 172; Mismatches 472; Indels 310; Gaps 42;

Qy 306 RVAPARDKSMVDPKKGTSLSDKL--SPSTHTLLGQFFQG----WGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLLFSLGCGIQRHCGKVLFLG 62
Qy 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHFQ--PFFRTNQVILTAPNR 417
Db 63 LLAFGALALGLRMAIETNLEQLWVEGSRVSEQLHYTKELGEEAAYTSQMLIQTAQOE 122
Qy 418 SSRYDSSLGLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQIDICY---AP 474
```


Db 123 GEN-----LTPALG-----LHQAALTSKQVQSVLSYKSWDLNKICYKSGVP 166
QY 475 LNPND-----TSLYDCCINSLOYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520
Db 167 LIENGMIENWIEKLPVCVILTPDCFWGAK---LQGSAYLPGRPDQIOWNLDPQELLE 223
QY 521 -----HFLYCANAPITFPKDGTAALASCHWADYA 548
Db 224 ELGPPFASLEGFRELIDKAQVQAYVGRPCLDHCPSPAPNHHSRQAPNVAHLSGCG 283
QY 549 PVPPP-----LAIGYKGDYSE---AEALIMTF-----SLNYPAGDPRLAQAKL 591
Db 284 HGFHKFMWQBELLGGMARPOGELLRAELQSTFLMSPRQLYEHFRGDIYTHDIGW 343
QY 592 WEEAFLEENRAFQRRMAGMFQVTTFAERSLE-DEINRTTAEDLPFATS-----Y 640
Db 344 SEEQASTVLQAWRRFVQLAQELPENASQOIHAFSSTLDDI-LHAFSEVAARVVG 402
QY 641 IVIFLYISALGSYSSRWVDSKATLGLGVAVVLGAVMAAGFFSYLGRSSVLILQ 700
Db 403 LMLAYACVTM---LRWD---CAQSGSVGLAGVLLVALAVASGLGICALLGITFNAATTQ 457
QY 701 VVPELVLSYGADNIRIFVLEYQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCPL 760
Db 458 VLFFLALGIGVODVFL--LAHAFTALPGTPIQERMBECLQRTGTGSVLTSINNMAAFILM 515
QY 761 GALTMPAVRTFALTSLGLAVILDFLQMSAFVALLSLOSKROEASRLDVCCC----- 812
Db 516 ALVLPALRAFSLQAAIVVGCTFVAVMLVFPAILSLDRLRRHCQELDLVLCFSPCSAQ 575
QY 813 ---VKPEL-----PPQOGEGL-- 827
Db 576 VIQILLPOELGDCGTPVPGIAHLTATVQAFTHCEASSQHVVTILPPQAHLVPPSPDLGSEL 635
QY 828 ---LIG-----PFQKAYAPFLLHMTTRGVLLFLA 855
Db 636 FSPGSTRDLGOEBEETROKACKSLPCARWNLAHFARYQFAPILLQSHAKAIVLVFGA 695
QY 856 LFGVLSYMSCHTSVGLDQBLAPKDSYLDYFLNRYFEVGAFFVFTVLGYNFS-SEA 914
Db 696 LGLSLYGATLVODGLATDVVPRGTKEHAFLSAQIRYFSL-YEVALVTQGFDAHNSQR 754
QY 915 GNAIACSSAGCNFFSFTQIKQVATFPPEOSYLAIPAS---SWDDFDIMLTPSSCC--- 967
Db 755 AL-----FDLHORFSSL-----KAVLPPEATQAPRTWLHYRYRWLQGIQAAFDQ 798
QY 968 ---RLYISGNKDKCFPSITVNSLNCILKNCMTGWSVPSVQFPHYLFWPLND 1018
Db 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGAQBPILD-FSOLATTRKLVLD 846
QY 1019 RENIKCPKGLAAYSTSVNLTSQGVLASRFWAYHKP----- 1055
Db 847 REGLIPPE--LFYMGITVWVSDPLGLAASQANFYPPPEWLHKDYDTTGEMLRIPPAQ 904
QY 1056 -----LKNSSQDYTEALRAARELAANI-TADLRKVPGTDPAPFEPVFTITNVFY 1102
Db 905 LEPAQFPFLRLGLQXTDPVEALEGARAACAEAGQVHAYPSGSPF-----LFW 954
QY 1103 BOYLILPEGLPMLSIC--LVPTFAVSCILLGLDIRSGLNLLSIVMLIVDTVGFMALWD 1160
Db 955 EOYGLG--RRCPLLAVCILLVCTFLVCALLLNPNWTAGLI-VLVLAMTMVELFGIMGFLG 1011
QY 1161 ISYNAVSLINYSVAGMSVEFVSHITRSPAISTKPTWLERAKETISMGSAVPAVAMTN 1220
Db 1012 IKLSAIPVILVASGVIGFEVTHVALGF-LTQGSRNLRHAHEHTFAPVTDG-AIST 1069
QY 1221 LPGAIVLGLAKAQLIQIPFFRLNMLLITLGLHLGLVFLPVILSVYVGPVDPNPALEOKRA 1280
Db 1070 LGLLMLAGSHDFIVRYFFAALTVLTLGLLHGLVLLPVLLSILGPP--PEVIQMYKES 1127
QY 1281 BEAVAAV-----MVASCNHPHSRVSTADNLYN 1308
Db 1128 PEILSPAPQGGGLRWGASSLIPQSFAVVTSMVTVAIH 1165

RESULT 30

US-09-293-505-7
; Sequence 7, Application US/09293505
; Patent No. 6348575

GENERAL INFORMATION:

; APPLICANT: de Sauvage, Frederic
; APPLICANT: Carpenter, David A.
; TITLE OF INVENTION: Patched-2
; FILE REFERENCE: P1405R1
; CURRENT APPLICATION NUMBER: US/09/293,505
; CURRENT FILING DATE: 1999-04-15
; EARLIER APPLICATION NUMBER: US 60/081,884
; EARLIER FILING DATE: 1998-04-15
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 7
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-293-505-7

Query Match 8.4%; Score 583.5; DB 2; Length 1182;

Best Local Similarity 24.2%; Pred. No. 2.3e-45;

Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;

QY 330 SPSTHTLLGQ-----PFOG-----WGTWVASWPLTILVLSVIPVVALAAGLVFTLT 376
Db 20 SSAPHILAGSLOAPLWLRAYFQGLLPSLCRIQKHGKVLFLGLVAFGALAGLVRVAVIE 79
QY 377 TDPVELMSPNSQARSEKAFHDQHFQ-PFFRTNQVILTAPNRSSYRSDSLLGPKNFSGI 435
Db 80 TDLEQLWVEVGRSVQELHYTKELGEBAAVTSQMLI---QTAHQEGGNVLTPE---A 131
QY 436 LDLDLLELLELQERLHLQWSPQAQNI SLODI CY---APLNPDN-----TSLYDCC 486
Db 132 LD-----LHQAALTSKQVQSVLSYKSWDLNKICYKSGVPLIENGMIERIEKLPFCV 184
QY 487 INSLILOYFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---FKDGTALA 539
Db 195 ILTPDLCFWGAK---LOGGSAYLPGRPDQIOWNLDPQELLELPFASLEGFRELIDKA 241
QY 540 LSCMADYGAPV-----PFLA---IGYKGDYS----- 565
Db 242 QVGQAYVGRPCLDPDHPCHPSPAPNRHSKQAPNVAQELSGGCHGFSHKFMWQBELLG 301
QY 566 -----BABALIMTF-----SLNYPAGDPRLAQAKLWBEAFLEENRAFQRRMAG 609
Db 302 TARDLQGLLRAELQSTFLMSPRQLYEHFRGDIYTHDIGW 361
QY 610 MFQVTFTAERSLE-----DEINRTTAEDLPFATSIV--IFLYISALGSYSSWS 658
Db 362 LAQELPANASQOIHAFSSTLDDI LRAFSE---VSTRVVGYYLLMLAYACVTMLRW 417
QY 659 RMVDSKATLGLGVAVVLGAVMAAGFFSYLGRSSVLILQVVPFVLVSVGADNIFIFV 718
Db 418 --CAQSGQAGVLAGVLLVALAVASGLGICALLGITFNAATTQVLPFALGIGVDDIFLLA 475
QY 719 LEYQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSG 778
Db 476 HAFTKAP--PDTPLPERMGECLRSTGTSTVALTSVNNMVAFMAALVPFIPALRAFSLQAAI 533
QY 779 AVILDPILLQMSAFVALLSLOSKROEASRLDVCCC-----VKQGE----- 817
Db 534 VVGCFNFAAMLVFPAILSLDRLRRHRQRLDLVLCFSPSPCSAQVIQMLPOELGDRVAVPGI 593
QY 818 ---LPP-----PG-----QGEQ-- 826
Db 594 AHLTATVQAFTHCEASSQHVVTILPPQAHLLSPASDPLGSELYSPGSTRDLSSEGTG 653
QY 827 -----LLLGFFQKAYAPFLLHMTTRGVVLLFLALFGVLSYMSCHISVGLDQ 873
Db 654 PQAACRPLCAHWTLAHFAFYQFAPILLQTRAKALVLLFFGALLGLSLYGATLVQDGLAL 713

QY 874 ELALPKDSYLDYFLFLNRYEFGAPVYVVTTLGYNFS-SEAGNNAICSSAGCANNFQ 932
DB 714 TDVPRGKHAFLSAQLRYFSL-YEVALVTGGGFDVAHSQAL-----FDLHQ 761
QY 933 KIQYATEPEQSYLAIPAS-----SWVDDF-----IDWLTSSCCRLYISGNK 976
DB 762 RFSSL-----KAVLPPPPATQAPRTWLHYRWSWLOGIQAFDQDWSGRITCHSYRNGSD 816
QY 977 D-----KFCFSTVNSLNCNKMSITM-----GSVRPSVEQFHKYLFWFLNDRPNKCPK 1026
DB 817 GALAYKLLIQTGNAQEPDLFQLTTRKLVDXEGLIPP--ELFYMGLTVMVSSDPL----- 869
QY 1027 GGLAA-----YST-----SVNLTSDQVLAASPMAYHKPLKNSQDYTEALR 1067
DB 870 -GLAASQANFYPPPELWHDKYDTTGENLRIPAAQPLEFAQFPFLHGLQKTADFVEAIE 928
QY 1068 AARELAANI-TADLRKVPCTDPAFEVFPYTTNVFYOYLTILPEGLFMLSIC--LVPTF 1124
DB 929 GARAACTEAGQAGVHAYPSGSPF-----LFWEQYGL--RRCFLLAVCILLVCTF 976
QY 1125 AVSCILLGLDLRSGLNLLSIWMLVDTVGFMAWMDISYNAVSLNLSAVGMSVEFVSH 1184
DB 977 LVCALLLSPWTAGLI-VLVLAWMTVEGLGFLGKLSAIPVVLVASIGIGVEFTVH 1035
QY 1185 ITRSPAISTKPTWLERAKATISMGSAVAGV---AMTNLPGLVILGLAKAQLQIIFFR 1241
DB 1036 VALGFLTSHGSRNRAA-----SALEQTFAPVTDGAVSTLLGLLMLAGSNFDFIIRYFV 1090
QY 1242 LNLTLTLLGLHGLFVLVSYGVP 1267
DB 1091 VLTVTLTLLGLHGLLVLVLSILGP 1116

RESULT 31
US-09-060-939A-7
; Sequence 7, Application US/09060939A
; Patent No. 6709838
; GENERAL INFORMATION:
; APPLICANT: Frederic de Sauvage, David A. Carpenter
; TITLE OF INVENTION: Patched-2
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; FILING DATE: 15-Apr-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Svoboda, Craig G.
; REGISTRATION NUMBER: 39,044
; REFERENCE/DOCKET NUMBER: P1405
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-1489
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1182 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-060-939A-7

Query Match 8.4% Score 583.5; DB 2; Length 1182;

Best Local Similarity 24.2%; Pred. No. 2.3e-45;
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;
QY 330 SFSTHTLLGQ-----PFQG-----WGTWASWPLTILVSVIPVVALAAGLVFELT 376
DB 20 SSAPHILAGSLQAPLMRLRAYFQGLFSLGCRIOKHCCKVLFLGLVAFGALALGLRVAVIE 79
QY 377 TDPVELWSAPNSQARSEKAFHDQHFQ-PPFRTNQVILTAPNRSSRYDSLLGLKPNFSGI 435
DB 80 TDLEQLMVEGSRVSQLHYTKELGEEAAYTSQMLI-----QTAHQEGGNVLTFE---A 131
QY 436 LDLDLLELLELEQLERHLQVWSPQEAONISLODICY---APLNPDN-----TSLYDCC 486
DB 132 LD-----LHLQAALTASKVQVSLYKSWDLNKICYKSGVPLTENGMIERMIKLPFCV 184
QY 487 INSLLQVFNQNRLLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT---FKDGTALA 539
DB 185 ILTPLDCFWEGAK---LQGGSAVLPGRPDIOQWNLDPQOLLEELGPFASLEGFRELIDKA 241
QY 540 LSCMADYGAVPF-----PFLA---IGYKGDYS----- 565
DB 242 QVGQAYVGRPCLDPPDPPHCPSPAPNRHSRQAPNVAQELSGGCHGFSHKFPMWQBELLLGG 301
QY 566 -----EAEALIMTF-----SLNNYPAGDPRLAQAKLWEEAFLEENRAFORRWAG 609
DB 302 TARDLQQLLAEALQSTFLLMSPRQLYEHFRGDYQTHDIGWSEEQASMWLQAWQRREVQ 361
QY 610 MFQVTFABRSLE-----DEINRTABDLIPATSYIV--IFYLISALGSYSWS 658
DB 362 LAQEALPANASQOIHAFSSTLDDILRAFSE---VSTTRVVGYYLLMLAYACVTMLRMD 417
QY 659 RMVDSKATGLGGVAVVLGAVMAAMGFSSVYLGIRSSLVILQVVPFLVLSVGADNIFV 718
DB 418 --CAQSQGAVGLAGVLLVALAVASGLGALLGTFNNAATQVLPFLALGIVGDDIFLLA 475
QY 719 LEYQRLPRPCEPREVHIGRALGRVAPSMLLCSSEAIKCFPLGALTMPAVRTFALTSG 778
DB 476 HAFTKAP--PDTPLPERMGECRLSTGTSTVLTNNMVAFFMAALVPTPALRAFSLOAAI 533
QY 779 AVILDFLLOMSAFVALLSLSKQEAERLDVCCC-----VKPQE----- 817
DB 534 VVGCNFAAVMLVFPAILSLDLRRHRORLDVLCFSSPCSAQVIQMLPQELGDRAPVGI 593
QY 818 -----LPP-----PG-----QGEQ-- 826
DB 594 AHLTRATVOAFTHCEASSQHVVTILPQAHLLSPASDPLGSELYSPGGSTRDLSQEGTG 653
QY 827 -----LLLGFFQKAYAPFLHLWITRGVLLLFLLALFQVSLYSMCHISVGLDQ 873
DB 654 PQAACRPILCAHWTLHAFYQFAPLLIQTAKALVLLFFGALLGLSLYGATLVQDGLAL 713
QY 874 ELALPKDSYLDYFLFLNRYEFGAPVYVVTTLGYNFS-SEAGNNAICSSAGCANNFQ 932
DB 714 TDVPRGKHAFLSAQLRYFSL-YEVALVTGGGFDVAHSQAL-----FDLHQ 761
QY 933 KIQYATEPEQSYLAIPAS-----SWVDDF-----IDWLTSSCCRLYISGNK 976
DB 762 RFSSL-----KAVLPPPPATQAPRTWLHYRWSWLOGIQAFDQDWSGRITCHSYRNGSD 816
QY 977 D-----KFCFSTVNSLNCNKMSITM-----GSVRPSVEQFHKYLFWFLNDRPNKCPK 1026
DB 817 GALAYKLLIQTGNAQEPDLFQLTTRKLVDXEGLIPP--ELFYMGLTVMVSSDPL----- 869
QY 1027 GGLAA-----YST-----SVNLTSDQVLAASPMAYHKPLKNSQDYTEALR 1067
DB 870 -GLAASQANFYPPPELWHDKYDTTGENLRIPAAQPLEFAQFPFLHGLQKTADFVEAIE 928
QY 1068 AARELAANI-TADLRKVPCTDPAFEVFPYTTNVFYOYLTILPEGLFMLSIC--LVPTF 1124
DB 929 GARAACTEAGQAGVHAYPSGSPF-----LFWEQYGL--RRCFLLAVCILLVCTF 976
QY 1125 AVSCILLGLDLRSGLNLLSIWMLVDTVGFMAWMDISYNAVSLNLSAVGMSVEFVSH 1184

Db 977 LVNALLLLSPWTAGLI-VLVLMAMTVELFGIMFGIKLSAIPVVLVASIGIVEFTWH 1035
QY 1185 ITRSPAISTKPTWLERAKEATISMGSAVPAGV---AMTNLPGLVILGLAKAQLIQIFFPR 1241
Db 1036 VALGFLTSHGSRNLRAA-----SALEQTPAPVTDGAVSTLLGLMLAGSNFPIIRYFPV 1090
QY 1242 LNNLITLLGLHGLVPLVILSVGP 1267
Db 1091 VLVTLTLLGLHGLLPLVLLSILGP 1116

RESULT 32
US-09-248-796A-16160
; Sequence 16160, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 16160
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-16160

Query Match 8.4%; Score 581.5; DB 2; Length 531;
Best Local Similarity 28.9%; Pred. No. 8.6e-46;
Matches 151; Conservative 100; Mismatches 220; Indels 51; Gaps 16;

QY 642 VIFLYISALGSSWSRVMVDSKATLGLGV-----AVVLGAVMAAMGFFSYLG 691
Db 13 IVFVGYYICF--FSPWQTTKQLVLIGENKIYIRVFENYNSIICNGV---SGIFSIIIG 67
QY 692 TRSSVILVQVPLVLSVADNIFTVLEYQRLPR-RGEPREVIHGRALGRVAPSMLLC 750
Db 68 LASTLIIAEVIFPLAIGIDNIFLIVHELHVISEGNPNLAEVRISQALKHIGSCFIS 127
QY 751 SLSEALCFPLGALTMPAARTFALTSGLAVIDLLOMSAFVALLSLDSKROEASRLDVC 810
Db 128 AVLQCMFLATSVGMPAVKNFAYYGAGAVLNFSLQMTCFGLLALDQRRLEDNRVDYV 187
QY 811 --CCVKP-----QBLPPPGQGGELLLGFFQKAYAPFLIHLHWTIRGVLILLFALPGVSIY 862
Db 188 PWTISPIQLQNDDEIDEPVHLEYNFSRWIGDHYAPFLKTKTKPKVITLVVLVWGISLS 247
QY 863 SMCHISVGLDQELALPKDSYLLDYFLNRYPEVAGPVVFTTLGYNFSSEKGMMAICSS 922
Db 248 LFPKIQGLDIDRIAIFSKSYLVNYSVVDYLNVPVFFVVK-DLDYSERLNQOKICGG 306
QY 923 -AGCNNEFTQKIQYATEPEQSYLAIPASSWVDPIWLTP--SSCCRLYISG--PNKD 977
Db 307 FSACDEFSLANTLEQEFKRSDISMSLESPASNLWDDFFSNLNPDLDOCCRFKKSTVFEKTP 366
QY 978 KFCPSVTNSLNLKNCMSITMGSRVPSVEQFHKILP-----WFLND---RPNKICPKGGL 1029
Db 367 EFCSPNAPORQC-QSCYL-----NHNPPYDSSMKAFPERDFMYFNDWIOQESDPCPLGGK 421
QY 1030 AAYTSVNLTSQGVLAGRFMAHYHKLKNSQDYTEALRAARELANITADLRKVPGETDPA 1089
Db 422 AAHQQAISRTTE-KIDSSYFRTSPAPLRGQDEFIYAYKSGNNIVKEITKLI-----PS 473
QY 1090 FEVPEYTTINVEYOYLITLPGFLWLSICLVPTFAVSCILL 1131
Db 474 MDVFAVSPFFIFFTQYQNVILLTVALLTVAMLIYIVISTFFL 515

RESULT 33
US-09-207-857-2
; Sequence 2, Application US/09207857
; Patent No. 6309879
; GENERAL INFORMATION:
; APPLICANT: Buncroft, David A.
; TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED
; TITLE OF INVENTION: THERETO
; FILE REFERENCE: ONV-05001
; CURRENT APPLICATION NUMBER: US/09/207,857
; CURRENT FILING DATE: 1998-12-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: human
US-09-207-857-2

Query Match 8.4%; Score 580; DB 2; Length 1203;
Best Local Similarity 22.9%; Pred. No. 5e-45;
Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLQQFPQG-----RGTWVASWPLTILVLS 359
Db 3 RSPPLRELPPSYTPPARTAAAPQLLAGSLKAPLWLRAYFQGLLFSLGGCQIRHCGKVLPLG 62
QY 360 VIPVVALAAGLVFTLTDPVELWSAPNSQARSEKAFHDQHG--PFFRTNQVILTAPNR 417
Db 63 LLAFGALGALGRMAIETNLEQLWVEGSRVSEQLHYTKELGEBAAVTSQMLIQIOTAROE 122
QY 418 SSVYDSLLGPKNFGSGLDLDLLELELELRHLQVMSPEAORNISLQDICY---AP 474
Db 123 GEN-----ILTPAIG-----LHLQAALTASKVQVSLYKSWDLNKKICYKSGVP 166
QY 475 LNPDN-----TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520
Db 167 LIENGMIERIEKLPCCVILTPDCFWEGAK---LQGSAYILPGRPDIOQNLDPQLLE 223
QY 521 -----HFLYCANAPLTFKDGFTALALSCMADYGA 548
Db 224 ELGPPASLEGFRELIDKAQVQVYGRPCLHPDDLHCPPSAPNHHRSQAPNVNHELSGC 283
QY 549 PVPPP-----LAIGYKGYKDYSE---AALIMTF-----SLNYPAGDPLRLAQA 591
Db 284 HGFSHKFMHQBELLGGMARDPOGELLRAELALQSTFLLMSPRQLYEHFRGDIYTHDIGW 343
QY 592 WEEAFLEEMRAPQRMMAGMFQVTFTAERSLE-DEINRTTAEDLPITFATS-----Y 640
Db 344 SEEQASTVLQAWQRFVQLAQEALPENASQQIHAFSSITLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISALGSSWSRVMVDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLVILQ 700
Db 403 LLMLAYACVTM--LRWD--CAQSGSVGLAGVLLVALAVASGLGLCALLGITTFNAATQ 457
QY 701 VVPFLVLSVGDADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFEL 760
Db 458 VLPFLAIGIGVDDVFL--LAHAFTALPGTFLQRMWEGECLORTGTSVVLTSINNMAFLM 515
QY 761 GALTMPAVRTFALTSGLAVIDLFLLOMSAFVALLSLDSKROEASRLDVC-- 812
Db 516 AALVPIPALRAFSLQAAIVVGCTFVAVMLVFPALFSLDLRREHQRDLVLCFSSPCSAQ 575
QY 813 ---VKPQL-----PPPGQEGEL-- 827
Db 576 VIQLPQBLGDTVPVGIHAHLTATVQAFTHCEASSQHVVTILPQOHLVPPSPDLGSEL 635
QY 828 -----LLG-----FPKAYAPELLHWTIRGVLILLPLA 855
Db 636 FSPGSGTRDILGQBEETKQKACKSLPCARWNLAHFARYQAPLLOSHAKAIVLVLPGA 695
QY 856 LFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGA PVYFVTLGYNFS-SEA 914


```
RESULT 35
US-09-807-007-1
; Sequence 1, Application US/09807007
; Patent No. 6881833
; GENERAL INFORMATION:
; APPLICANT: ZAPHIROPOULOS, Peter et al.
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALING PATHWAY
; FILE REFERENCE: 2921-0130P
; CURRENT APPLICATION NUMBER: US/09/807,007
; CURRENT FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 1146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-007-1

Query Match      8.4%; Score 579; DB 2; Length 1146;
Best Local Similarity 23.3%; Pred. No. 5.7e-45;
Matches 276; Conservative 161; Mismatches 452; Indels 298; Gaps 40;

QY 306 RVAPARDKSMVDPKKGTSLSDKL--SFSTHTLLGQFFQG---WGTWVASWPLTILVLS 359
DB 3 RSPPLRELPSVTPPARTAAPQILAGSLKAPLWLRAYFQGLLFSGLCGIQRHCGKVLFLG 62
QY 360 VIPVVALAAGLVFTELTPDVELMSAPNSQASEKAFDQHFQ--PPFTNQVILTPNPR 417
DB 63 LIAFGALAGLRAIETNLEQLWVEVGRSVQELHYTKELGEEAAYTSQMLIQTPAQE 122
QY 418 SSRYDLSLGLPKPNSGILDLLELLELQERLRLHLOVSPSAQRNLSLDICV---AP 474
DB 123 GEN-----ILTEALG-----LHQAALTASKVQVSLYKGSWDLNICKYKSGVP 166
QY 475 LMPDN-----TSLYDCCINSLLQVPPNNRTLLLTANQTLMGQTSQVQDWK----- 520
DB 167 LIENGMIERMIKLPFCVILTPDCFWGAK---LQGSAYLPGRPDIQWTNLDPEQLIE 223
QY 521 -----HFLYCANAPLTFKDGITGALALSCWADYCA 548
DB 224 ELGPPASLEGFRELIDKAVQGYVGRPCPLHPDDLHCPPSPAPNHHSRQAPNVABELSGGC 283
QY 549 PVPPFP-----LAIGYKGDYSE---AEALIMTF-----SLNNYPAGDPRLAAQAKL 591
DB 284 HGFSHKFMWQBELLLGGWAPDQOELLRAELAQSTFLLMSRQLYEHFRGDIYTHDIGW 343
QY 592 WEEAFLEENRAFQRMAGMFQVTTAERSLE-DEINRTTAEDLPIFATS-----Y 640
DB 344 SEEQASTVLQAWRRFPVQLAQEALPENASQIHAFFSSTLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISALGSYSWSRWVDSKATIGLGAVVVLGAVNMAAGFFSVLGRSSLVILQ 700
DB 403 LLMLAYACVTM---LRWD---CAQSGSVGLAGVLLVALAVASGLGICALLGITFNAATQ 457
QY 701 VVPFLVLSGADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEALCPPL 760
DB 458 VLFPALGIGVDVDFL---LAHAFTALPGTLPQERMGECIQTGTSVLTSINNMAFLM 515
QY 761 GALTMPAVRTPALTSGLAVIDLFLQMSAFVALLSLDSKQEASRLDVCC----- 812
DB 516 AALVFPALRAFSLQAAIVVGCTFFVAVMLVFPAILSLDLRRRHCRQLDLVLCFSSPCSAQ 575
QY 813 ---VPPQEL-----PPGQGEGL-- 827
DB 576 VTQILPQELGDTVPVGIHAHLTATVQAFTHCEASSQHVVTILPPOAHLVPPSPDPLGSEL 635
QY 828 -----LLG-----FFQKAYAPELLHMITRGGVLLPLA 855
DB 636 FSPGSGTRDLQGEETROKAAKCKSLPCARWNLAHFARYQFAPLLLSQAKAIVLVLPGA 695
QY 856 LFGVLSYNSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGAPVYFVTTILGYNFS-SEA 914
```

```
DB 696 LLGLSLYGATLVQDGLALTDVVPRGTHKHAFLSAQLRYFSL-YEVALVTQGGFDYAHQSQR 754
QY 915 GMAICSSAGCNFSPFTQKIQYATEPPEQSYLAIPAS-----SWDDFDIDLWLTSSCC--- 967
DB 755 AL-----FDLHQRFSSL-----KAVLPPPPATQAPRTWLHYRNWLOGIOAAFDQ 798
QY 968 -----RLYISGNKDKFCPCSTVNSLCKNCKMSITWGSVRPSVQFPHKYLWFLND 1018
DB 799 DWASGRITRHSYRG-SED-----GALAYKLLIQTGDAQELLD-FSOLTTTRKLVLD 846
QY 1019 RPNIKCPKGLAAYSTSVNLTSDGQVLAISRFMAYHKP----- 1055
DB 847 REGLIPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEPEMLHDKYDTTGENFRIPPAQP 904
QY 1056 -----LKNSSQDYTEALRAARELAANI-TADLRKVCTDPAFVFPFTTINVFY 1102
DB 905 LEPAPFPFLRLGLQKTADFEAIEGARAAACABAGQAGVHAYPSGSPF-----LPW 954
QY 1103 EOYLTLPEGLFMLSIC--LVPTFAVSCILLGLDLRSGLNLLLSIVMLLVDTVGFMAWLD 1160
DB 955 EOYGLG--RRCFLAVCILLVCTFLVCALLLNPNWTAGLI-VVLAMTVFELFGMGFLG 1011
QY 1161 ISYNAVSLINLVSAYGMSYFVSHITRSPAIKPTMLERAKEATISMGSAVPAGVAMTN 1220
DB 1012 IKLSAIPVIVLASVIGVETVHVALGF-LTQGSRLRAAAHLEHTFAPVTDG-AIST 1069
QY 1221 LQGLVLGLAKAQLIQIPFRNLNLLITLGLHLGLVFLVILSYGCP 1267
DB 1070 LGLGLLAGSHPDFIVRYFFAALTVLTLLGLHLGLVFLVILSILGP 1116

RESULT 36
US-08-540-406-4
; Sequence 4, Application US/08540406
; Patent No. 5837538
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-540-406-4
```

```
Query Match 7.6%; Score 526.5; DB 1; Length 1311;
Best Local Similarity 22.2%; Pred. No. 7.1e-40;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LLGQFFQCGWGTWASWPLTILVLSVIVPVALAAGLVFTLTTPDVELWSAPNSQARSEKA 395
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 ILGCFQGGDAGKVLFP--VAIIVLSFTCV-----GLKSAQHTRVDQLWVQEGGRLEALK 124
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 396 FHDQHFQGF-PRTNQVILTAPNRSSRYRSDLSLLGPKNFSGLDLDLLELLEQLERLHL 454
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 YTAQALGEADSSHTQLVI---QTAQKPDVSLHP-----GAL-----LEHL 162
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 455 QWMSPEAQRI-----SLODICVAPLNPD-----NTSLYD-----CCINSLLQYFQNN 497
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 163 KVVHAATRVTVHMYDIEWRLKDCYSPSPDPFEGYHHIESIDNVIPCAIITPLDCFWEG 222
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 498 RTLL-----LITANQTLMGOTSQ--- 515
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 223 SKLGFDPDPIVYVPHLKHKLQWTHLNPLEVVEEVKCLKFQPLSTIEATMKRAGITSATMK 282
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 516 ---VDWKDHFLYC-ANAPLTFKDGDTALALSCMADYGAPVFPF-----LAIGG----- 558
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 283 KCLDPTDP--HCPATAP-NKSGHI PDVAABELSHGCGYFAAAVYHWPQELIVGATRNS 339
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 559 ----YKGDYSAEALIMTFSLNNTYPADGPRLAQAALWEEAFLEEMRAPORFEMAGMFQVT 614
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 340 TSALRKARXLQTVVQLMGEREMYWADHYKHQIGWQEKAAAVLDWQRKFAAEVRKI 399
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 615 FT-----AERLEDEINRTAEDLPIPATSVIVIFLYISLALGYSYSSWSRV 661
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 400 TTSGSVSSAYSYPSTSTLNDILGKFSEVLSKXNIIIGYMFMLIVAVTL---IQW-RDP 455
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 662 VDSKATGLGGVAVVLGAAMAAGFFSYLGIRSSILVILQVVPFLVLSVGCADNIFIFVLEY 721
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 456 IRSQAGVGIAGVLLSITVAAGLGFCAIGIPFNASSQTIVFFLALGLGVQDMFLTHTY 515
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 722 QBLRRPGPREVHIGRALGRVAPSMILCSLSEATCFIFGALTMPAVRTFALTSGLAVI 781
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 516 --VEQAGDVEERTGLVLKSGLSVLLASLNCVMAFLAAALLPIPAFVFCLOAAIILL 573
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 782 LQFLLOMSAFVALLSDSKROASRLDVCCVKPQELPP----- 820
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 574 FNIGSILLVFPAMISLDRRRGAARDLCCCLMPSPLPKKIPIERAKTRKNDKTHRIDT 633
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 821 -----PGQEGGL-----LLGFFQKAYAPFLHHTITRGVLLLFLLALFGVLSYMC 865
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 634 TRQPLDPDVSENVTKTCLSVSLTKWAKQYAPFIMRPAVKVTSMLALIAVILTSVWGAT 693
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 866 HISVGLDQELALPKDSYLDYFLFNR---YFVVGAPVYFVTTLGVNFSSEAGMNAICSS 922
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 694 KVKDGLDLTDIVPENT---DEHEFLSRQEKYGF-----YN-----MYAVTQ- 732
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 923 AGCNFSPF--TKIQYATFPEQSYLAIP-----ASSWVDQFIDWL----- 961
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 733 ---GNFEYPTNQKLLY--EYHDQ-FVRIPIIKNVNGGLTKFWLSLFRDRLDQLQVAPDK 786
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 962 -TPSSCCRLYISGPNKDKFCPTSVNSLNCIKNCMSITMGSVRPSVSEQ----- 1007
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 787 EVASGCI-----TQEYMKCNASDE-GILAYKLMVQTGHVDNPDIKSLITAGHRLVDKD 838
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1008 -----FKHYL-PWFNDRPNKCPKGGI-----AAYSTSVNLTSDGOVLASRFM 1050
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQFORWIHSPEDVHLEIKKSSPLIYITQLP 898
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1051 AVHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITITNVFVEQYITILP 1110
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 899 FYLSGLSDTYSIKITLRSVRDLCLKVEA--KGLPN-----FPGIPLFWEQYLYLRT 949
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1111 EGLFMISLCIVPTFAVSCLLGLDLRLSGLNLSIWMILVDVTGVFMALMDISYNAVSLIN 1170
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 950 SLLALACALAAVF-IAMVYLLNMAAAVLVTLATLVLQLLGVNALLGVKLSAMPVIL 1008
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


QY 1282 EAVAVMVASCPNHPRSVST 1301
DB 1111 -----EHPERLST 1118

RESULT 42
US-09-754-032-4
; Sequence 4, Application US/09754032
; Patent No. 6921646
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; JOHNSON, RONALD L
; ADDRESS: FOUR EMBARCADERO CENTER, SUITE 3400
; CITY: SAN FRANCISCO
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-754-032-4

Query Match 7.6%; Score 526.5; DB 2; Length 1311;
Best Local Similarity 22.2%; Pred. No. 7.1e-40;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;

QY 336 LIGQFFQGGTGVASWPLTILVLSVPPVVALAAGLVFTLTDPVELASPNQSRSEKA 395
DB 72 ILGCLQGDAGKVLFP--VAILVLSFTCV-----GLKSAQIHTRVDQWVQEGGRLEALK 124

QY 396 FHDQHFQGP-FTNQVILTAPNRSYRYSLLGPNFSGILDLDLLELLELQRLRLH 454
DB 125 YTAQALGEADSTHQLVI-----QTAQDPVSLHF-----GAL-----LEHL 162

QY 455 QWSPQEAQRNI-----SLQDICYAPLNPDP-----NTSLYD-----CCINSLLQYQNN 497
DB 163 KVVHAATRVTHMYDIEWRLKOLCYSPSPDPFEGYHHIESIIDNVIPCAIITPLDCFWEG 222

QY 498 RTLL-----LITANQTLWGQTSQ--- 515

DB 223 SKLLGPDPIYVPHLKHKLQWHLNPLEVVEEVKLLKQFPLSTTEATMKRAGITSYMK 282

QY 516 ---VDMKHFLYC-ANAEPLTFKDGFTALALSCWADYGAFFPPF-----LAIGG----- 558

DB 283 KPCLDPTDP--HCPATAP--NKKSCHIPDVAABLSHGCGYFAAAYMHMPQLVGGATRNS 339
QY 559 ---YKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLSENRFAFORMAGNFQVT 614
DB 340 TSALKRARKLQTVVQJMGREMYEYADHYKHVQHWGNOEKAAVADWQVFAAEVRKI 399
QY 615 FT-----AERSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSYSSSRVM 661
DB 400 TTSGSVSSAYSPPFSTSTLNDILGKFSEVSLKNILGYMFLIYVAVTL---IQW-RDP 455
QY 662 VDSKATLGLGCVAVVGLVAMAMGPFYSYLGIRSSILVILQVPPFLVSVGADNIFIVLSY 721
DB 456 IRSQAGVGIAGVLLLSITVAAGLFCALIGIPFNASTQIVPFLGALGLGVQDMFLITHTY 515
QY 722 QRLPRPGEPREVIHGRALGRVAPSMLLCSSEALCFPLGALTTPMPAVKTEALTSGLAVI 781
DB 516 --VEQAGDVPREERTGLVKKSGLSVLASLCVNAFLAAALLPFPAPFVCLQAAAILLL 573
QY 782 LDFLQMSAFVALLSLDSKQBEASLDVCCCVKQBELPP----- 820
DB 574 FNLGSILLVPPAMISLDLRRRSARADLLCCLMPESPPLPKKI PERAKTRKNDKTHRIDT 633
QY 821 -----PGQEGCL-----LLGFFQKAYAPFLHMTITRGVLLLLFLALFGVSYSMC 865
DB 634 TRQPLDPDVSENVTKTCLSVSLTKWAKNQYAPFIMRPAVKVTSMLALIAVILTSVMGAT 693
QY 866 HISVGLDQELALPKDSYLLDYFLNLR---YEVGAPVYVFTVLGYNFSSBMAAICSS 922
DB 694 KYDGLDLDIVPENT---DEHEFLSRQEKYPGF-----YN-----MYAVTO- 732
QY 923 AGCNMFSE--TQKIQYATEFPEQSYLAIP-----ASSWVDDFDIDL- 961
DB 733 ---GNFEYPTNQKLY--EYHDQ--FVRIPNIIKNDNGGLTKFWLSIFRDLMLDLQVAPDK 786
QY 962 -TPSSCCRLYISGPNKDKFCPSTVNSLNCNMCNITMGSVRPSVEQ----- 1007
DB 787 EVASGCI-----TQEWCKNASDE-GILAYKLMVOTGHVDNPDKSLITAGHRLVDKD 838
QY 1008 -----FHKYL--PMFLNDRPNIKCPKGL-----AAYSTSVNLTSDQVLSRPM 1050
DB 839 GIINPKAFNYLSAWATNDALAYGASQGNLKPQORWIHSPEDVHLEIKKSSPLIYTQLP 898
QY 1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITINVFYEQLTILP 1110
DB 899 FYLSGLSDTXSIKTLIRSVDLCLEYA--KGLPN-----FPSGIPFLFWEQIYLYLT 949
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLSRGLNLLSIVMLVDTVGFPMALWDISYNVSLIN 1170
DB 950 SILLALACALAAVF-IAVMVLLNAAVAVLTALATLVQLGLGVNALLGVKLSAMPAVL 1008
QY 1171 LVSAGMSVEFVSHITRSFALSTKPTWLERAKEATISMGSAVPACVAMTNLPGILVLGLA 1230
DB 1009 LVLAIGRGVHPTVHLCLGFTVSIGC-----KRRASLALES-VLAPV-----VHGALAAALA 1059
QY 1231 KAQL-----IQTPFFRLNLLITLLGLLHGLVFLPVLISYVGP--DYNPALALRQKRAE 1281
DB 1060 ASMLAASECGFVARLFLRLDDIVPLGIDGLFPFVLSILGPAAEVRPI----- 1110
QY 1282 EAVAVMVASCPNHPRSVST 1301
DB 1111 -----EHPERLST 1118

RESULT 43
US-08-916-140-4
; Sequence 4, Application US/08916140
; Patent No. 6946257
; GENERAL INFORMATION:
; APPLICANT: Scott Mathew P.
; APPLICANT: Goodrich, Lisa V.
; APPLICANT: Johnson, Ronald L.
; APPLICANT: Epstein, Ervin Jr.

; TITLE OF INVENTION: PATCHED GENES AND USES RELATED THERETO	
; FILE REFERENCE: CBST-904-203	
; CURRENT APPLICATION NUMBER: US/08/916.140	
; CURRENT FILING DATE: 1997-08-21	
; PRIOR APPLICATION NUMBER: US 08/656,055	
; PRIOR FILING DATE: 1996-05-31	
; PRIOR APPLICATION NUMBER: US 08/540,406	
; PRIOR FILING DATE: 1995-10-06	
; PRIOR APPLICATION NUMBER: US 08/319,745	
; PRIOR FILING DATE: 1994-10-07	
; NUMBER OF SEQ ID NOS: 64	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 4	
; LENGTH: 1311	
; TYPE: PRT	
; ORGANISM: Butterfly	
; FEATURE:	
; NAME/KEY: VARIANT	
; LOCATION: 348, 908	
; OTHER INFORMATION: xaa = Any Amino Acid	
US-08-916-140-4	
Query Match	
Best Local Similarity 7.6%; Score 526.5; DB 2; Length 1311;	
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;	
QY	336 ILGQFFQGMGTWASWPLTILVLSIPVVALAAGLVFTLTDPVELMSAPNSQARSEKA 395
DB	72 ILGFCLOGDAGKLVF--VAILVSTFCV-----GLKSAQIHTRVDQVMWQEGGRLEAEK 124
QY	396 FHDQHFQGF-FTNQVILTAPNRRSYRYSLLGPNFSGILDLLLELLELQERLRL 454
DB	125 YTAQALGEADSSHTOLVI---QTAKOPDVSLLHP---GAL-----LEHL 162
QY	455 QWSPQAQRNI-----SLODICYAPLNPD-----NTSLVD---CCINSLLQYFQNN 497
DB	163 KVHHAATRVTHMYDIEWRLKDLCSPIPDPEGYHHIESIDNVIPCALITPLDCFWEG 222
QY	498 RTLL-----LTANQTLMGTSQ--- 515
DB	223 SKLLGPDYPIYVPHLKHKLQWTHLNPLEVVEVKLKFQPLSTIEAYMKRAGITSAYMK 282
QY	516 ----VDWKDFLYC-ANAPLTFKDGITALALSCHADYCAPVPPF-----LAIGG----- 558
DB	283 KPCLDPTDP--HCPATAP-NKSGHIPDVAABELSHGCGYGAFAAAYMHPPEQLIVGGATRNS 339
QY	559 ----YKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEENRPAQRMAGMFOVT 614
DB	340 TSAIRKARXLQTVVQJMGREMYEYWDHYKHQIGWQEKAAAVLDADWQKFAAEVRKI 399
QY	615 FT-----AERLEDEINRTTAEDLPFATSYIVIVFLYISLALGSYSWSRV 661
DB	400 TTSGSVSSAYSFYPPFSTSLNDILGKFSVSLKNILGYMFLIYVAVTL---IQW-RDP 455
QY	662 VDSKATLGGVAVVGVAAVMAAGPFSYIGIRSSVLIVQVFPFLVSGADNIFIFVLEY 721
DB	456 IRSQAGVGIAGVLLSITVAAGLGFCAILGIPFNASSTQIVPFLALGLGVQDMFLTHY 515
QY	722 QRLPRPGEPREVHIGRALGRVAPSMLLGSLSEACIFFLGAITPMPAVRTFALTSLGLVI 781
DB	516 --VEQAGDVPREERTGLVYKXGSLSVLLASLNCVMAFLAALLPIPAFRVCLQAILLL 573
QY	782 LDFLLQMSAFVALLSDSKRQEAASRLDVCCVKPQELPP----- 820
DB	574 FNLSILLVFPAMISLDRRSARADLLCCLMPESPFPKKKIPERAKTRKNDKTHRIDT 633
QY	821 -----PCQEGE-----LLGFQKAYAPFLHWTIRGVVLLFLALFVGLSYSMC 865
DB	634 TROPLODPDVSNTYKTCCLSVSLTKWAKNQYAFIMRPAVKVTSMALIAVILTSVWGAT 693
QY	866 HISVGLDQELAKPKDSYLLDYFLFLNR---YFEVGAPVYFVTTVLGYNFSSEAGMNAICSS 922
DB	594 KVKQGLDLTDIVPENT---DEHEFLSKQKYFGF-----YN-----MYAVTQ- 732
PCT-US95-13233-4	
PCT-US95-13233-4	
; Sequence 4, Application PC/TUS9513233	
; GENERAL INFORMATION:	
; APPLICANT: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY	
; TITLE OF INVENTION: Patched Genes and their Use	
; NUMBER OF SEQUENCES: 19	
; CORRESPONDENCE ADDRESS:	
; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert	
; STREET: Four Embarcadero Center, Suite 3400	
; CITY: San Francisco	
; STATE: CA	
; COUNTRY: US	
; ZIP: 94111	
; COMPUTER READABLE FORM:	
; MEDIUM TYPE: Floppy disk	
; COMPUTER: IBM PC compatible	
; OPERATING SYSTEM: PC-DOS/MS-DOS	
; SOFTWARE: PatentIn Release #1.0, Version #1.30	
; CURRENT APPLICATION DATA:	
; APPLICATION NUMBER: PCT/US95/13233	
; FILING DATE: 06-OCT-1990	
; CLASSIFICATION:	
; ATTORNEY/AGENT INFORMATION:	
; NAME: Rowland, Bertram I	
; REGISTRATION NUMBER: 20015	
; REFERENCE/DOCKET NUMBER: a60190-1	
; TELECOMMUNICATION INFORMATION:	
; TELEPHONE: 415-781-1989	
; TELEFAX: 415-398-3249	
; INFORMATION FOR SEQ ID NO: 4:	
; SEQUENCE CHARACTERISTICS:	
; LENGTH: 1311 amino acids	
; TYPE: amino acid	
; STRANDEDNESS: single	
; TOPOLOGY: linear	
; MOLECULE TYPE: protein	
PCT-US95-13233-4	

Query Match	7.6%	Score 526.5	DB 4	Length 1311
Best Local Similarity	22.2%	Pred. No. 7.1e-40		
Matches 258	Conservative 175	Mismatches 420	Indels 307	Gaps 46
Qy	336	LLGQFFGCGWGWASWPLTTLVLVSIPVVALAAGLVFTELTDPVELWSAPNSQARSEKA	395	
Db	72	ILGCFLOGDAGKVL P--VALVLSTFCV-----GLKSAQHTRVDQLWVGEGGLEAEKL	124	
Qy	396	FHDQHGFPP--FRTNQVILTAPNRSSRYRSDLSLLGPKNFGSGLDLDLLELLELQERLHL	454	
Db	125	YTAQALGEADSSITHQLVI---QTAQDPDVSLLHP---GAL-----LEHL	162	
Qy	455	QVNSPEAQRNI-----SLQDICYAPLNDP-----NTSLYD-----CCINSLLOYFQNN	497	
Db	163	KVHAATRVTVHMYDIEWRLKDLCSYSPIDPFGYHHIESIIDNVICAITPLDCEFWEG	222	
Qy	498	RTLL-----	515	
Db	223	SKLLGPPYPIYVPHLKHKLQWHLNPLEVVEEVKLGKQPPSLSTIEAYMKRAGITSAYMK	282	
Qy	516	---VDWKDHPFLYC--ANAPLTPFKOCTALALSCMADYGAPVFPF-----LAIGG-----	558	
Db	283	KPCLDPTDP--HCPATAP--NKSGSHIPDVAEELSHGCGFAAAYMHWPQLIVGGATENS	339	
Qy	559	----YKGDYSEAEALIMTSLNNYPAGDPRLAQAQKLWEBAFLIEMPAFORMMAGMPQVT	614	
Db	340	TSALRKARXLQTVVQLMGEREMEYADHYKHQIGMNOEKAAAVLDAWQKFAAEVRKI	399	
Qy	615	FT-----AERSLEDEINRTAEDLPFPATSYIVIFLYISLAGSVSSWSRVM	661	
Db	400	TTSGSVSSAYSFPFPSTSTLNDILGKSESVSLKNIILGYMFMLIYVAVTL---IQW-RDP	455	
Qy	662	VDSKATLGLGSAVVLGAVNAAMGFFSYLSIRSLVLQVVPFLVLISVGADNIPFVLEY	721	
Db	456	IRSOAGVGIAGVLLSITVAAGLGFALLGIPFNASSTQIVPFLALGVLGVDMPLLTHTY	515	
Qy	722	QRIPRRPGEPREHVHIGALGRVAPSMLLCSLSBAICFFLGALTMPAPVTRFALTSLGLAVI	781	
Db	516	--VEQAGDVPREERTGLVLKKSGLSVLLASLCNMAFLAAALPAPAPRVFCLOAAILL	573	
Qy	782	LDPELLQMSAFVALLSLDSKQESRLDVCVCVQDELPP-----	820	
Db	574	FNLGSIILVFPAMISLDLRRRSARADLLCCLMPESPLPKKIPERAKTRKNDKTHRIDT	633	
Qy	821	-----PGQEGEL-----LILGFPQAYAPFLHLWITRGVLLLLFLALFGVLSYMC	865	
Db	634	TRQPLDPDVSENVTKTCLSVLTKWAKQYAPFIMRPAPKVSTMLALIAVILTSVWGAT	693	
Qy	866	HISVGLDQELALPKOSYLLDYFLPLNR---YFEVGAPVYFVTTLGYNFSSEAGMNAICSS	922	
Db	694	KVKDGLDLTDIVPENT--DEHEFLSRQKYFGF-----YN-----MYAVTQ--	732	
Qy	923	AGCNPFSP--TKIQYATEFPQSYLAIP-----ASSWDDPIDWL-----	961	
Db	733	---GNPEYPTNKKLLY--EYHQD--FVRIPNIIRKNDNGGLTKFMLSFLRDLQLQVAFDK	786	
Qy	962	TPSSCCRLYISGNKDKPFCSTVNSLNCCLKMCSITMGSVRSVEQ-----	1007	
Db	787	EVASGCI-----TQYWCNKNASDE--GILAYKLWVQTGHVDNPIDKSLITAGHRLVDDK	838	
Qy	1008	-----FKHYL--PWFILNDRPNIKCPKGGI-----AAYSTSVNLTSDDGOVLASRFM	1050	
Db	839	GIINPKAFNYLSAWATNDALAYGASQNLKPKQPMQRMHSPEDVHLBEIKKSSPLIYVQLP	898	
Qy	1051	AYHKPLKNSODYTEALBAARELANITADLRKVPQGTDPAREVFPYTTINVFYQYLTLIP	1110	
Db	899	FYLSGLSDTXSIKTLIRSVRDCLIKYEA--KGLPN-----FPSGPIFLWFYQYLYLRT	949	
Qy	1111	EGLFMLSCLVPPFAVSCLLGLDLRSGLLNLISIVMILVDVTGFWALMDISYNVNSLIN	1170	
Db	950	SLLLALACALAAVF--IAVMVLLNANAVLVLTALATVLVQLGVWALLGVKUSAMPVIL	1008	
Qy	1171	LVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNPGLVILGLA	1230	

Db	1009	LVTALGRGVHTVHLCLCGFVTSIGC-----KRRRASLALES-VLAPV-----VHGALAAALA	1059
Qy	1231	KAQL-----IQIFPERLNLITLLGLHGLPLPVLISVGP--DVNPALALEQKRAE	1281
Db	1060	ASMLAASECGFVARLFURLGLDIVFLGLIDGLLFFPVLVSILGPAAEVRPI-----	1110
Qy	1282	EAVAAVMVASCPNHPSPRVST	1301
Db	1111	-----EHPERLST	1118
RESULT 45			
US-09-268-140-3			
; Sequence 3, Application US/09268140			
; Patent No. 6268176			
; GENERAL INFORMATION:			
; APPLICANT: Gemmill, Robert M.			
; TITLE OF INVENTION: Trachin, Robert A.			
; FILE REFERENCE: 93445-00004			
; CURRENT APPLICATION NUMBER: US/09/268,140			
; CURRENT FILING DATE: 2000-03-12			
; PRIOR APPLICATION NUMBER: US 60/077,723			
; PRIOR FILING DATE: 1998-03-12			
; NUMBER OF SEQ ID NOS: 46			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 3			
; LENGTH: 1286			
; TYPE: PRT			
; ORGANISM: Drosophila melanogaster			
US-09-268-140-3			
Query Match 6.9%; Score 474; DB 2; Length 1286;			
Best Local Similarity 20.6%; Pred. No. 6.6e-35;			
Matches 246; Conservative 176; Mismatches 465; Indels 306; Gaps 43;			
Qy	321	KGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVSVIPVVALAAGLVFTTELTTDPV	380
Db	43	KGKARGSRTAIYLRSPVQSHLETIGSSVQKHGKVLFAILVLSTFCVGLKSAQIHSKVH	102
Qy	381	ELMSAPNSQARSKAFHDQHFQ--PFRTNQVILTA--PNRSSRYVDSLLLGKPNFSGIL	436
Db	103	QLMIQEGRLAEALAYTKTIGDEBATHQLLIQTTHDPNASVLHPQALL-----A	153
Qy	437	DLDLLELLELQERLRLQVWSPQAQNTLSQICVAPLNPNTSLY-----DCC	486
Db	154	HLEVLVKATAVKVHLVDTE-W-----GLRDMCMNPSTPSPEGIYYIEQILRHLP	204
Qy	487	INSLQYFQN-----NRTLLLTAN-----OTLMGQTSQVDWKDH	521
Db	205	IITPLDCFEGSQLLGPESAVVIPGLNORLLWTLNPNASVMQTMKQNSEEKISPDFETV	264
Qy	522	FLY-----C-----ANAPLTFKD-----GTALALSCWADYGAPVPFP	553
Db	265	EQWKMRAATGSGYMEKPCPLNPNCPDTPAKNKTQPPDVGAILSGCYG-YAAKHMW	323
Qy	554	---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----	590
Db	324	PBELIVGGRKRNRSGHLRKAQALQSVQVLMTKEKMYDQMDNVKVVHLGWTQEKAAEVLN	383
Qy	591	LWEEAP---LEEMRAQRQRMAGHFQVFTFAERSLEDEINRTAEDLPFATSVIVIFLYI	647
Db	384	AWQRNFSREVEQLLRKQSRITATNYDIVVSSAALDDILAKFHPSPALSIVIGVAVTVLY-	442
Qy	648	SLALGYSYSSWRVMVDSKATPLGLGVAVVLGAWMAAMGFFSYLIGIRSSIVILQVVPFLVL	707
Db	443	--AFCTLLRW-RDPVRGQSSVGAVGLVLCMCFSTAAGLGLSALLGIVFNAASTQVVPFLAL	499
Qy	708	SVGADNIFIVLBYQRLPRRPGPREVHIGRALGRVAPNMLCSISEATCFPLGALTMP	767
Db	500	GLGVYDHIFMLTAAYASBNRR--EQTKL-----ILKKVGPISILFSACSTAGSFFAAFIPVP	553

994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKG 1028
813 ILAYKLIYQTHVDNPDVKELVNLNRLVNSDGIINORAFYNLSAW-----ATNASSPTEL 868
1029 LAA--YSTSVNLTSDGOVLASRFMYHKLKNSQD-----YTBALRAAR 1070
869 LRANCIRNRANGASQGLPYPRQYFHP--NEYDLKPKSLPLVYAQMPPYLHGLDTS 926
1071 ELAANITADLRKVPCTDPAFV--PPYITNVFVEQYLTILPEGLFMLSCLVTFPAVSC 1128
927 QI-KTLIGHIRDLSVKYEGFGLNVPYGPPIFEQYMT-LRSSLAMILACVLLAALV 984
1129 LLLGLDRLSGLNLLSIYMLVDTVGFMAWDISYNAVSLINLVSACMSVEFVSHITRS 1188
985 SLLLSVAAVLVLSLASLAQIFGAMTLLGKLSAIPAVILISVGMMLCF--NVLIS 1042
1189 FAISTKPTWLRKAKETISMSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRNLNLT 1248
1043 LGFTSVGNRRRVOLSMQSLGPLVHGMLTSGVAVFMLSTSPFEFVIRHFCWLLLVLC 1102
1249 LGLLHGLVFLPVLSYVGPVNPALALEOKRAEAAVAAVMVASCNHPSRVSTADNIYV 1308
1103 VGACNSLLVFPILLSMWGPAB-LVPLE-----HPDRISTSPPLVR 1143
1309 HS 1310
1144 SS 1145
RESULT 48
US-08-954-698-48
Sequence 48, Application US/08954698
Patent No. 6271363
GENERAL INFORMATION:
APPLICANT: Ingham, Phillip W.
APPLICANT: McMahon, Andrew P.
APPLICANT: Tabin, Clifford J.
TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
TITLE OF INVENTION: Proteins and Uses Related Thereto
NUMBER OF SEQUENCES: 48
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY, HOAG & ELIOT LLP
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109-2170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,698
FILING DATE: 20-OCT-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/462,386
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/435,093
FILING DATE: 04-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/356,060
FILING DATE: 14-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/176,427
FILING DATE: 30-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: HMV-006.10
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000

TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 1299 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-954-698-48
Query Match 6.7%; Score 464.5; DB 2; Length 1299;
Best Local Similarity 20.4%; Pred. No. 5.4e-34;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
321 KGTSLSDKLSFSTHTLIGQFFQGWTVWASWPTILVLSVIPVVAALAGLVFTLTDPV 380
43 KGKARGSRTAIYRSVFQSHLETIGSSVQKHAGKVLFAVLVLTFCVGLKSAQIHSHKVH 102
381 ELMSAPNSQARSEKAFHQHFG-PFFRTNQVILTA-----PNRSSYRYSLLLGPKNFSGIL 436
103 QLMIOEGGGLEAELAVTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
437 DLDLLELLELQERLHLQVWSPQAQRNLSLQDIQYAPLNPNTSLY-----DCC 486
154 HLEVLRKATAVKVHLXYDTE-W-----GLRDMCNMPSTSPFEGIYIIEQILRHLPCS 204
487 INSLQYFQN-----NRTLILLTAN-----QTLMGQTSQVDMKH 521
205 IITPLDCFWESQLLGPESAVVIGLNQRLWLTLPASVMQYKQMSBEKISFDPTV 264
522 FLYCANAPLTPKOGTALALSCM-----ADYGAPVFPPLAIG--GYK 560
265 EQYMKRAAIA--SGYMEKPCLNPLNCPDTPAKNKSTOPODVG-----ILSGCYGYA 317
561 GKDYSEBALIMTFSLNYPAGDPLAQA-----589
318 AKHMHPEELIVGAKGN-RSGHLRKAQALQSVVQLMTEKEMQYQWQNDYKVHGLWTQE 376
590 -----KLWERAF--LEEMRAFQRMAGMQVFTTAERSLEDEINRTABDLPIFATSY 640
377 KAAEVLANWQNRFSREVEQLLRKQRIATNYDIVVFSSAALDDILAKSPHSALSIVIGV 436
641 IVIFLYISLALGSSYSSWSRVNDSKATLGLGGVAVLGAVMAMGFFSYIGRSLVLIQ 700
437 AVTVLY--AFCTLLRW-RDPVRGQSSVGAVLLMCFSTAAGLSALLGVFVNAATQ 492
701 VVPLVLSVGADNTFIFVLEYQRLPRRGPGEPRVHIGRALGRVAPSMLLSLSEAIQFPL 760
493 VVPLALGLGVHDHIFMLTAAYAESNR--EQTKL-----ILKVGPSILFSACSTAGSFFA 546
761 GALTMPAVRTFALTSGIAVILDFLLQMSAFVALLSLDSKROEASRLDV-CCCV-----K 814
547 AAFIPVPALKVFCLOAAIVMCSNLAAALLVFPAMISLDRRTAGRADIFCCCFPWVWKEQ 606
815 PQLEPP-----PQCEGL-----LLGFF 832
607 PKVAPPVPLPANNNGRGARPKSCNNNRVALPAQNPLEQRADIFGSSHSLSASFSLATFA 666
833 QKAYAPFLLHWTTRGVVLLLLFLALFGVSLYSNMCHI SVGLDQELALPKDSYLLDYFLFLNR 892
667 FQHYTFPLMRSWVKFLTVMGFLAALISSIYASTRIQDGLDIIIDLVPKDSNEHKFLDAQTR 726
893 YFEVGAPVYVTTGLGYNFSSSEAGMNAICSSACNNPFSFTKIQYATEPPEQSYLAIP----- 949
727 LF--GFYSMYAVTQG-----NFEVPTQQLLRDY-HDSFVRVPHVI 764
950 -----ASSWVDDFDIWLTPSSCCRLYISGNKDKFCPSTVNSLNLKNC-----993
765 KNDNGGLPDFMLLFSWL-----GNLQKIFDEEYDRDGLTKCECFPPNASSDA 812
994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKG 1028
813 ILAYKLIYQTHVDNPDVKELVNLNRLVNSDGIINORAFYNLSAW-----ATNASSPTEL 868

QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQILQIPFFRLNLLITL 1248
Db 1043 LGFMTSVGNRQRVOLSMQMSLGLPLVHGLMTSGVAVFMLSTSPFVFIRHFCWLLLVLC 1102
QY 1249 LGLHGLVLPVLTSVVGPDVNPALALEQKRAEEAAVAVMVASCPNHPBSRVSTADNIYN 1308
Db 1103 VGCNSELVFPILLSMVGPEAE-LVPLE-----HPDRISTPSPLPVR 1143
QY 1309 HS 1310
Db 1144 SS 1145

RESULT 51
US-08-954-128-48
; Sequence 48, Application US/08954128
; Patent No. 6610656
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; APPLICANT: McMahon, Andrew P.
; APPLICANT: Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; TITLE OF INVENTION: Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,128
; FILING DATE: 20-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/462,386
; FILING DATE: 05-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 04-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-128-48

Query Match 6.7%; Score 464.5; DB 2; Length 1299;
Best Local Similarity 20.4%; Pred. No. 5.4e-34;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDKLSPFTHLLGQFGQGWASWPLTLLVLVSVIPVVAAGLVFTLTDVDP 380
Db 43 KGKARGSRTAIYLSRVFQSHLETAGSVQKHAGKVLVFAILLVSTFCVGLKSAQIHSKVH 102

QY 381 ELMSAPNSOARSEKAFHQHFG-PFRTNQVILTA---PNRSSRYDLSLLGPKNFSGIL 436
Db 103 QLWIQSGGGLAEALATVQKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELELERLHLQVWSPQAORNISLODICVAPLNPNTSY------DCC 486
Db 154 HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPSTPSPEGIYYIEQILRHLIPCS 204
QY 487 INSLLOQYFQ-----NRTLILLITAN-----OTLMQOTSQVDWKDH 521
Db 205 IITPLDCFEGSQQLGPESAVVIPGLNQRLLTTLNPASVMQMKMQMSEKISFDFTV 264
QY 522 FLYCANAPLTPKDGTAALASC-----ADYGAPVPPFLAIG--GYK 560
Db 265 EQMKRAAIA---SGWMEKPCLNPLNCPDTPAPKNSTQPPDVGA---ILSGCGCYGA 317
QY 561 GKDYSEAEALIMTFPSLNNYPAGDPRLAQA-----589
Db 318 AKHMHMPEELIYGGAKRN-RSGHLRKAQALQSUVQLMTEKEMYDQDNKYKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFORRMAGMFOVPTTAERSLEDEINRTTAEDLPATSY 640
Db 377 KAAEVUNAWQRPNSREVEQLLRQSRIATNYDIYVSSAALDDILAKFSHPSSALSIVIGV 436
QY 641 IVIFLYISLALGYSYSWSRVMDSKATLGLGGVAVVVLGAVMAAMGFFSYLGRSSSLVILQ 700
Db 437 AVTVLY---APCTLLRW-RDPVRGQSSVGVAGVLLMCPSTAAGLGLSALLGIIVFNAATQ 492
QY 701 VYFPLVLSVGNADNIFIVLEYQRLPRPCGEPREHIGRALGRVAPSKLCSSEATCFLL 760
Db 493 VYFPLALGLGVDHIFMLTAAYAESNRR--EOTKL-----ILKKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTTALTSGLAVILDFLQMSAFVALLSLDSKROEASLDV-CCCV-----K 814
Db 547 AAPIPVPALKVFLQAAIVMCSNLAAALLVFPAMISLDLRRRTAGRADIPCCCFPWKQE 606
QY 815 PQLELP-----ASSWVDDFIDMLTPSSCCRLYISGPNKDKFCPTVNSLNCNKNC-----993
Db 607 PKVAPPVPLNNGNGARHPKSCNNRVALPAQNPLLEQRADIPSSHSLSFSLATPA 666
QY 833 QKAYAPFLHWTIRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLDYFYFLNR 892
Db 667 FQHYTPFLMRSWVKFLTVMGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVVTTLGVNPSSEAGMAICSSAGCNFSPQKIQYATFEPEQSVIATP---949
Db 727 LF--GFYSMYAVTQG-----NFEYPTQOQLLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWVDDFIDMLTPSSCCRLYISGPNKDKFCPTVNSLNCNKNC-----993
Db 765 KNDNGGLDPFWLLLPSEWL-----GNLQKIPDEEYRDGLTKECFPNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PMFLNDRPNIKPKGG 1028
Db 813 ILAYKLIQVQGHVDPVDKELVLNRLVNSDGIINQRAFNYLSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVNLTSDGOVLASRFMAYHKPLNSQD-----YTEALRAAR 1070
Db 869 LEANCIRNRANGASQGLYPEPRQYFHPQ--NEYDLKIPKSLPLVYAQMPFYLHGLTDT 936
QY 1071 ELAANITADLRKVPGTDPAFEV--PPYTTITNVPYQYLTILPEGLFMLSCLCVPTFAVSC 1128
Db 927 QI-KTLIGHIRDLVKYEGFGLPNTPSGIPFPFIMQYMT-LRSSLAMILACVLAALVLV 984
QY 1129 LLLGLDLRSGLNLLSIVMILVDTVGFMAIMDISYNAVSLINLNAVSGMSVBFVSHITRS 1188
Db 985 SLLLSVMAAVILSVLASLAQIFGAMTLLIGIKLSAIPAVILILSVGMWLCF--NVLIS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQILQIPFFRLNLLITL 1248
Db 1043 LGFMTSVGNRQRVOLSMQMSLGLPLVHGLMTSGVAVFMLSTSPFVFIRHFCWLLLVLC 1102

Db	154	HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPESTPSFEGIYIYIEQLRHLP	204
QY	487	INSLQYFON-----NRTLILLTAN-----QTLMGQTSQVDWKDH	521
Db	205	IITPLDCFWEGSOLLGPESAVVIPGLNQRLLTINPASVMQYMKQKMBEIKSFDFTV	264
QY	522	FLY-----C-----ANAPLTFKD-----GTALALSCMADYGAPVPPF	553
Db	265	EQYMKRAAIGSGYMEKPCPLNPLNCPDPAFNKNSQPPDVGAILSGGCYG-YAAKHMHW	323
QY	554	---LAIGGVK---GKDYGEAEALIMTFSL-----NNYPAGDPRLAQAK-----	590
Db	324	PEELVGRKRNRSRHLRKAQALQSVQLMTEKEMYDQNDQNYKVHHLGWTQERAAEVLN	383
QY	591	LWEEAF---LEEMRAFQRRMAGMFOVTTAERSLEDEINRTAEDLPFATSYIVIFLYI	647
Db	384	AWQRNFSREVEQLLRKQSRATNYDIYVFSAAALDDILAKFSHPALSIVIGVAVTVLY-	442
QY	648	SLALGSYSNRSVMVDSKATLGLGGVAVVLGAVMAAMGFFSYLGI-----	692
Db	443	--AFCTLLRW-RDPVRGGOSSVGAVGLLMCFSTAAGLSALLGIVFNALTAAAYAESNR	499
QY	693	-RSSIVI---LQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRALGRVAPSM	747
Db	500	EQTLLKNAQVPPFLALGVDHIFI-----VGFSI	533
QY	748	LCSLSEATCFPLGALTMPAVRTFALTSGLAVIDFLQMGAFVALLSLDSKQAEARL	807
Db	534	LFSACSTAGSFFAAAFIPVPAKLVFCLQAAIWMCSNLAALLVFPAMISLDLRRRTAGRA	593
QY	808	DV-CCCV-----KPOELPP-----PQ	823
Db	594	DIFCCCFPVWKQOPKVPVPLPLNNGRGARHPKSCNNRVLPAPQNPLLQORADIPGS	653
QY	824	GEGL-----LLGFQKAYAPFLHWTIRGVVLLFLALFGVSLYSCHISVGLDQBLALPK	879
Db	654	SHSLASFSLATFAQHYYTFPLRSWVKFTVMGFLAALISSIYASTRLQDGLDIIDLVPK	713
QY	880	DSYLLDYFLFNRYFVGVAPVFTVTLGYNFSSEAGMNAICSSAGCNFSPFKQIYATE	939
Db	714	DSNEHKFLDAQTRLF--GFYSYVAVTQG-----NFEYPTQOQLLRD	752
QY	940	FPEQSVLAP-----ASSWVDDFDMLTPSSCCRLYISGPNKDKFCFSTVNSLNC	989
Db	753	YHDS--FRVPHVTKNDNGGLPDMFLLLFSEWL-----GNLQKIFDEEYRDGRL	798
QY	990	LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF	1015
Db	799	TKECWFPNASSDAILAYKLVQTHGVDPVDRKELVLTNRLVNSDGIINQRAFNVYLSAWA	858
QY	1016	LNDRPNIKCPKGLAAYSTSVNITSGQVLASRFMAYHKLKNSQD-----	1061
Db	859	TND-----VFAVG-----ASQGLYPEPROVPHOP--NEYDKIPKSLPLVYAQM	901
QY	1062	--YTEALRAARELAANITADLRKVPGCTDPAFEV--PPYITVNYFXYQVLTILPEGLFMLS	1117
Db	902	PFLHGLTDSQI-KTLIGHIRDLSVKYEGFLNTPSGIPPIFWEQYMT-LRSSLAMIL	959
QY	1118	LCVLPTFAVSCLLIGLDSGLNLLSIYMLIVDTVGFMAIWDISYNVSLINLSAVGM	1177
Db	960	ACVLLAALVLVLSLLLSVWAAVLVLSLASLAQIFGAMTLLGILKLSAIPAVILLSVGM	1019
QY	1178	SVFVSHITRSPAISTKPTWLERAKZATISMGSAVFAGVAMTNLPGLVGLAKAQI	1237
Db	1020	MLCF--NVLLISLGFMITSVGNRRQSVOLSMQMSIGPLVHGMLTSGVAVFMLSTSPF	1077
QY	1238	FFRFLNLLITLLGLHGLFPLVLSVGVDPVNPALALEQKRAEEAAVAVMVASCNHP	1297
Db	1078	HFCWLLLVLCVGCNLSLVFPILLSWVGPEAE-LVPLE-----HPD	1118
QY	1298	RVSTADNIYVNH	1310

Db

1119

RISTPSPLPVRSS

1131

RESULT 59

US-09-754-032-6

; Sequence 6, Application US/09754032

; Patent No. 6921646

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/754,032

; FILING DATE: 03-Jan-2001

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/540,406

; FILING DATE: 06-OCT-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1285 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; SEQUENCE DESCRIPTION: SEQ ID NO: 6:

US-09-754-032-6

Query Match

Best Local Similarity

Matches 242;

Conservative 174;

Mismatches 450;

Indels 347;

Gaps 44;

QY

321

KGTSLSDKLSPSTHTLLQFFQCGWTWASWPLTILVLSVIPVVALAAGLVTELTDPV

380

Db

43

KGKARGRTAIYLRVSFQSHLETGLSSVQKGAGKVLFAILLVLSTFCVGLKSAQIH

102

QY

381

ELWAPNSQARSEKAFHQHFG-PFFRTNQVILTA---PNRSSRYDLSLLGPKNPSGIL

436

Db

103

QWIOEGGRLEALAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A

153

QY

437

DLDLLELLELLEQLRHLRHLQVMSPEAQRNISLQDICYAPLNPNSTLSY-----

486

Db

154

HLEVLVKATAVKVHLYDTE-W-----GLRDMCMPESTPSFEGIYIYIEQLRHLP

204

QY

487

INSLQYFON-----NRTLILLTAN-----QTLMGQTSQVDWKDH

521

Db

205

IITPLDCFWEGSOLLGPESAVVIPGLNQRLLTINPASVMQYMKQKMBEIKSFDFTV

264

QY

522

FLY-----C-----ANAPLTFKD-----GTALALSCMADYGAPVPPF

553

Db

265

EQYMKRAAIGSGYMEKPCPLNPLNCPDPAFNKNSQPPDVGAILSGGCYG-YAAKHMHW

323

Qy	554	---	LAIGYK	--	GKDYSEAEALIMTFSL	-----	NNYPAGDGRLLAQAK	-----	590																																															
Db	324	PEELIVG	GRKRN	RGHLKA	QALOSV	VQLOAT	KEKMYDQWQNYKVHGLWTQ	QBAAEVLN	383																																															
Qy	591	LWEEAF	---	LEEMRA	FORMAGM	FOVTF	TAERSLE	DEINRTTAEDLP	FIATSYIVILFI	647																																														
Db	384	AWQNF	SEVQLL	RKQRI	ATNYD	IYVPS	AAUDD	ILAKF	SHPSALS	IVIGVATVLY	442																																													
Qy	648	SLALGS	YSWSR	VMVDS	KATLGL	GGVAV	VLGVAN	MAAMG	FFFSYLG	-----	692																																													
Db	443	--	AFCTL	LRW	-RDP	VRGQS	VGVA	GLLMCP	STAGLS	ALLGIVFNAL	TAAYAESNR	499																																												
Qy	693	-RSSIV	----	LOVVP	FLVSV	GADNT	FI	FVLE	QRLPRR	CEPRE	VHIGRALGR	VAPSM	747																																											
Db	500	EQTLL	IKNA	STQV	VPFL	ALGL	GV	DHFI	-----	VGPSI	533																																													
Qy	748	LLCSL	SEAI	CFPL	GAL	TPMA	VRT	FALT	SLGL	AVILD	FLLOMS	AFVALL	SLDS	KROE	SRL	807																																								
Db	534	LFSAC	STAG	SFAA	PT	VP	PA	LK	VFC	LQA	AI	VMCS	NLAAL	VL	VP	AMIS	LDLRR	TAGRA	593																																					
Qy	808	DV	-CCCV	-----	KQELPP	-----	PGQ	823																																																
Db	594	DI	FCCP	PPV	MEQ	PKV	APP	VL	PLNN	NGR	GAR	HPK	S	CNN	NR	VP	LP	PAQ	NLL	EQ	RA	D	PGS	653																																
Qy	824	GEG	L	---	LG	FO	K	A	V	A	P	L	L	H	W	I	T	G	V	L	L	L	F	A	L	F	G	V	S	L	S	M	C	H	I	S	V	G	L	D	O	E	L	A	P	K	879									
Db	654	SHSL	AS	FL	AT	F	A	F	Q	H	T	P	P	L	M	R	S	W	K	V	F	L	T	G	M	G	F	L	A	L	I	S	S	T	A	S	T	R	L	D	I	I	D	L	V	P	K	713								
Qy	880	DSYLL	D	F	L	N	R	V	E	G	A	P	V	V	T	T	L	G	N	P	S	S	E	A	G	N	A	I	C	S	S	A	G	C	N	N	P	S	T	O	K	I	O	V	A	T	E	939								
Db	714	DSN	E	K	F	L	D	A	G	T	R	L	P	-	G	F	Y	S	M	A	V	T	O	-----	NPEY	T	O	Q	Q	L	L	R	D	752																						
Qy	940	P	PEO	S	Y	L	A	I	P	-----	ASSW	V	D	F	I	D	W	L	T	P	S	C	C	R	L	Y	I	S	G	N	K	D	K	F	C	P	S	T	V	N	S	L	N	C	989											
Db	753	YHDS	-	FRV	P	H	V	I	K	N	D	G	L	P	D	F	L	L	L	F	S	E	M	-----	GN	L	K	I	F	D	E	Y	R	D	G	R	L	798																		
Qy	990	LKNC	-----	MSIT	G	S	V	R	P	S	E	Q	-----	FHK	Y	L	-PWF	1015																																						
Db	799	TKC	W	P	P	N	A	S	S	D	A	I	L	A	K	L	I	V	O	T	G	H	V	D	N	P	V	D	K	E	L	V	L	T	N	R	L	V	N	S	D	G	I	I	N	R	A	F	Y	N	L	S	A	W	A	858
Qy	1016	LND	R	P	N	I	K	C	P	K	G	L	A	A	S	T	S	V	N	L	T	S	D	G	O	V	L	A	S	R	F	M	A	V	H	K	P	L	K	N	S	D	-----	1061												
Db	859	TND	-----	VFA	G	-----	ASQ	K	L	P	E	P	R	Q	Y	F	H	Q	P	-----	NEYD	L	K	I	P	K	S	L	P	L	V	A	Q	M	901																					
Qy	1062	--	YTE	A	R	A	E	L	A	N	T	A	D	L	R	K	P	G	T	D	P	A	F	E	V	-----	P	F	Y	T	T	N	V	F	E	O	Y	L	T	I	P	E	G	L	P	M	L	S	1117							
Db	902	P	F	Y	L	H	G	L	T	D	T	S	Q	I	-	K	T	L	I	G	H	R	D	L	S	V	K	E	G	F	L	P	N	P	S	G	I	P	F	I	P	E	Q	Y	M	T	-	L	A</							

RESULT 60

US-08-916-140-6

; Sequence 6, Application US/08916140

; Patent No. 6946257

GENERAL INFORMATION:

APPLICANT: Scott Mathew P.

APPLICANT: Goodrich, Lisa V.

APPLICANT: Johnson, Ronald L.

Db 714 DSNEHKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752
Qy 940 PFEQSYLAIP-----ASSWVDDFIDMLTSSCCRLVSGNKKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDGGLPDFMLLFSEWL-----GNLQKIFDEEYRDGRL 798
Qy 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLI VQTGHVDPVKELVLTNRLVNSDGI INQRAFYNLSAWA 858
Qy 1016 LNDRENIKCPKGLAAYSTVNLTSDDGOVLASRFMAYHKPLKNSOD----- 1061
Db 859 TND-----VFAYG-----ASQKLYPEPRQYFHOP--NEYDLAKI PKSLPLVYAQM 901
Qy 1062 --YTALRAARELANITADLRKVPCTDPAPEV--FPYITNVFEQYLTILPEGLFMLS 1117
Db 902 PFYHGLTDTSQI-KTLIGHIRDLVKYEGFGLPNYPGIPPIFEWQYMT-LRSSLAMIL 959
Qy 1118 LCLVPTFAVSCILGLDLRSGLNLLSVIILVDTVGFMAWDISYNAVSLINLVSAGM 1177
Db 960 ACVLLAALVSLLLSVAAVLVILSVLASAQIFGAWTLGIGKLSAIPAVILSVGM 1019
Qy 1178 SVEFVSHITRSPAISTKTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQLIQI 1237
Db 1020 MLCF--NVLLISLGFMTSGNRRVQLSNQMSLGLPLVHGMLTSGVAVFMLSTSPPEFVIR 1077
Qy 1238 PFRNLNLTLLGLHLVFLPVILSYGPOVNPALALEQKRAEAAVAVMVASCPNHP 1297
Db 1078 HFCWLLVVLVCAGCNLLVFPILLSMVGPEAE-LVPLE-----HPD 1118
Qy 1298 RYSTADNIVNHS 1310
Db 1119 RISTSPLPVRSS 1131

RESULT 61

PCT-US95-13233-6

; Sequence 6, Application PC/TUS9513233

; GENERAL INFORMATION:

; APPLICANT: THE BOARD OF TRUSTEES OF TEH LELAND STANFORD JUNIOR UNIVERSITY

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US95/13233

; FILING DATE: 06-OCT-1990

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Rowland, Bertram I

; REGISTRATION NUMBER: 20015

; REFERENCE/DOCKET NUMBER: a60190-1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-781-1989

; TELEFAX: 415-398-3249

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1285 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US95-13233-6

Query Match 6.3%; Score 438.5; DB 4; Length 1285;
Best Local Similarity 20.0%; Pred. No. 1.6e-31;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;
Qy 321 KGTSLSKLSPSTHTLLGQFFQGTWVWASWPLTILVLSPVVALAAGLVFTELTTDPV 380
Db 43 KKGARGSRTAIIYLSVFSQSHLETGLSSVQKHAGKVLFAILLVSTFCVGLKSAQIHSKVH 102
Qy 381 ELWAPNSQARSEKAFHQHFG-PFRTNQVILTA---PNRSSYRYSLLILGPKNFSIL 436
Db 103 QLWIQEGRLAEALAYTKTIGEDESATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DLDLLLELLEQLERLHLQVWSPAQRNISLQDICVAPLNDPNTSLY-----DCC 486
Db 154 HLEVIVKATAVVHLYDTE-W-----GLRDCWNPSTFEGIYIEQILRLHLIFCS 204
Qy 487 INSLLQYFQN-----NRTLLLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFWEGSQLLPESAIVIPGLNORLLWTLNLPASVMQYMKQKMSSEKISDPFETV 264
Qy 522 FLY-----C-----ANAPLTFKD-----GTALALSCMADYAGPVPPF 553
Db 265 EQYMKRAAIGSYMEKPCINLPNCPDPAFNKNSSTOPDVGAILSGCYG-YAAKHMH 323
Qy 554 ---LAIGCYK---GKOYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db 324 PEELIVGGRKNRSGHLRKAQLOSVOVLTEKEMYDQWQNYKVHHLGWTQEKAAEVLN 383
Qy 591 LWEEAF---LEEMAFQRMAGMFQVTFEASRLEDEINRTTAEIDLPIFATSYIVIFYI 647
Db 384 AWQNFSEVQELARKQSRIATNYDIYVFSSAALDDILAKFSPHSALSIVIGVAVTVLY- 442
Qy 648 SLALGSYSSWSRVWVDSKATGLGCVAVVLGAVNMAAGFFSYLGI----- 692
Db 443 --AFCTLLRW-RDPVRGQSSVGAVGLMCFSTAAGLSALLGIVFNALTAAAYAESNRR 499
Qy 693 -RSSIIVI---LQVVPFLVLSVGADNIFI FVLEYQRLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVPFALGLGVDDHIFI-----VGRSI 533
Qy 748 LLCSLSEACFPFLGALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSLSKRQKQASRL 807
Db 534 LFSACSTAGSFFAAAFIPVPALKVFCLOAAIVMCSNLAAALLVPFAMISLDLRRRTAGRA 593
Qy 808 DV-CCCV-----KPOELPP-----PQ 823
Db 594 DIFCCCFPVWKEQPKVAPFVPLNANNNGRGARHPKSCNNRVLPFPAQNPPLLEQRADIPGS 653
Qy 824 BEGL-----LLGFFOKAYAPFLLHWITRGVLLFLFALFCVSLYSMCHISVGLDQELALPK 879
Db 654 SHSLASFSLATFAFOHYTFPLMRSWVKFLTWGFLAALISLSYASTRLQDGLDIIDLVPK 713
Qy 880 DSYLLDYFLNRYPEVGAPVYVTTGLNVPSSSEAGMNAICSSAGCNPFSTQKIQYATE 939
Db 714 DSNEHKFLDAQTRLF--GFYSMYAVTQG-----NFEYPTQQQLLRD 752
Qy 940 PFEQSYLAIP-----ASSWVDDFIDMLTSSCCRLVSGNKKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDGGLPDFMLLFSEWL-----GNLQKIFDEEYRDGRL 798
Qy 990 LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db 799 TKECWFPNASSDAILAYKLI VQTGHVDPVKELVLTNRLVNSDGI INQRAFYNLSAWA 858
Qy 1016 LNDRENIKCPKGLAAYSTVNLTSDDGOVLASRFMAYHKPLKNSOD----- 1061
Db 859 TND-----VFAYG-----ASQKLYPEPRQYFHOP--NEYDLAKI PKSLPLVYAQM 901
Qy 1062 --YTALRAARELANITADLRKVPCTDPAPEV--FPYITNVFEQYLTILPEGLFMLS 1117
Db 902 PFYHGLTDTSQI-KTLIGHIRDLVKYEGFGLPNYPGIPPIFEWQYMT-LRSSLAMIL 959

US-09-332-522E-6

Query Match 2.2%; Score 150; DB 2; Length 1237;
Best Local Similarity 19.3%; Pred. No. 0.00037;
Matches 120; Conservative 86; Mismatches 240; Indels 176; Gaps 21;
QY 340 FFOGHTWVA-----SWPL-----TILVLSVIPWALAAAGLVTELT----- 377
DB 10 FSKAWFPACGCHGTFPGYFLINIPGTFITKLVVPEYFGSGSLSHSLNTSTTPQE 69
QY 378 -----DPVELWSPNSQARSEKAFHQHFGFPFRTNQVIL---TAPNRSRYDS 424
DB 70 PHSGEPWPBPQVLSNTTDRSPPLLPWAQSSPAFFVQOITLRTSVLEPWEQMLMD 129
QY 425 LLLGPNFSGILDLDLLELELQRLHLQWSPQAORNISLQDICYAPLNPONTSLYD 484
DB 130 AFRAP-----LHEVFKLLEIVRNHQ--SSENKR--TLEHNC--LHVDNVK--R 169
QY 485 CCINSLLQYFQNNRTLLLLTANOTLMGQTSQVDWKDHFYCANAPLTFKDGTAALASCM 544
DB 170 GTHGQLDQIFBYGCLLSPAN-----LWQNSQNFTRDTNILN----- 208
QY 545 DYGAFFVFPFLAIGYKGDYSAEALIMTFSLNYPAG---DPLAQAKLMEAEF----- 596
DB 209 -----TIFQVHNLQKSKVSAEMLFGLPMQDTGFKRYPLKARSRIIQYALTFLK 258
QY 597 ---LEEMRAFQRMAGMFQVTTAERSLEDEINRT-----TAEDLPFATSIVIVFL 645
DB 259 HNDMEYLDLTKELKLRHPPLPLASASABEPTTITIFYPGSEYRMMELVPYTVAPMLVFA 318
QY 646 YISALGVSYSRSMVW--DSKATLGLGVAVVLGAVMAAGFYSYLGIRSSILVILQVVPF 704
DB 319 YV-----YFSVRKIDVFRSRLALCSVITTAGSLANSLGCLFPFGTISLQSKDIFPY 372
QY 705 LVLSVGADNIFIVLEYQRLPRRPGEPREVHTGRALRVAPSMMLCSLSEAICFFLGALT 764
DB 373 LVILVLENSLIVTKSVVSMDET--FDVKIRVAQALSKEGWHISKTLTTEITILTIGLAT 430
QY 765 PMPAVRTFALTSGLAVIDFLQMSAFVALLSLSKQOE-----GLLLG----- 830
DB 431 FVEVQEFCEIFAIVGLLSDFMQLLFTILAMNKRTEYTABAKHLPMLLSCCTQAGR 490
QY 807 LDVCCVKPQELPP--PGQGE-----GLLLG----- 830
DB 491 QDRFREGAAPALPPFVGTGFSQSHPKLCPADPASVSDTSILVNGHSSPEQRIPIKIKI 550
QY 831 -----PFQAYAPFLHWI 844
DB 551 VNFWARTRFQAFMIWMIVMI 572

RESULT 65

US-09-252-991A-17657
; Sequence 17657, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17657
; LENGTH: 1467
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17657

Query Match 2.1%; Score 145; DB 2; Length 1467;
Best Local Similarity 19.8%; Pred. No. 0.0015;
Matches 269; Conservative 157; Mismatches 444; Indels 488; Gaps 71;
QY 176 RVPAAATLAVGTGCVYGSALCNQRMWLNFGDTGNGLAPLDITFHLLEPQAVSGI-- 233
DB 303 RQGRATDAGAVRQOQAGDR--QASR-----AGAGRRP-----FHRHRPGOALCAGRRR 349
QY 234 -----QPLNEGVARCNESQG-----DDVATCS-----CQD 258
DB 350 RQDPVPAGTRPDGRRPAGRPAGRAHRRGPGPAGHADHTAPCGNRWHARRPYED 409
QY 259 CAASCPAIARPOALDSTFYLGQMEGSLVLIILCSFAVV--TILLVFRVAPADKSKM 316
DB 410 RGRHTMTFPR-----FFIDR-----PIFAIVLSVLMIGGIVS----- 442
QY 317 VDPKKGTSLSKLSFSTHTLLGQFFQ--GWGTWVASWELTILVLSVIP-----VVALAAGL 370
DB 443 -----PFQPLSEYPAVTPPTVQVTTAYPGANPDVIAQTAT 479
QY 371 VFTELTTPVELMSAPNSQARSEKAF-----HDQHFQPF-----RTNOVILTAPNR- 417
DB 480 PLEQAITG--VEGMLYMSQSATDGRMILTITFDQIHDPDMAQIQVQNRVSRVLSLPEV 538
QY 418 -----SSRYDLSLLGPKNFGILDLDLLELLELQERL----- 451
DB 539 QROGVVTKTSPDILMVVHLLSPQRYDPLVI--SNVAYLQVRDELLRLPGISDVVWGA 596
QY 452 --RHLOW--SPE--AQRNLSQDICYAPLN-----PDNTSLYDCCINSL--L 491
DB 597 GEYSRMLWLDPLIAARGLTAGEVIAAVREQNVQVAGAVGQAPDSTAFOVTNTLGR 656
QY 492 QYFQNNRTLLLTANOTLMGQTSQVDWKDHFYCANAPLTFKDGTAALASCMADYGA 551
DB 657 SDEQFQDIIVRTGAD---GVVTRL-----RDVARIEMGGDAYALRSLLD--GEAV 703
QY 552 PFLAI---GGYKGDYSEAEALIMTFSLNYPAGDPLAQAKLMEAEFLEEMRAFQRM 608
DB 704 ALHIIQSPGAY--ALDTAESVRATVARLEGNFPAG-----LSARIA 742
QY 609 GMFQVTFTAERSLEDEINRTTAEDLPFATSIVIVFLYISALGYSYSWSRVWDSKAT- 667
DB 743 --YDPTVFVRASLQTV--ATTLLERAILLVVVVWLFL-----RSWRASLIPIPMV 789
QY 668 LGLGVAVVLGAVMAAGFYSYLGIRSSILVILQVVPFLVLSVG--ADNIFIFVLEYQRLP 725
DB 790 VSLVGTFF---AVHLMGF-----SLNTLSLFG--LVLSIGIVVDDAIVVVENVERHI 836
QY 726 RRPGEPREVHIGRALGRV--APSMMLCSLSEAICF---FLGALTMPAVRTFALTSG 781
DB 837 EN-GEPPLOQAARRAMDEVTPGIMATTSVLAADVFIPTAFLSLGLOG--EFYRQFALT 894
QY 782 LDFLLQMSAFVALLSLSKQREASRLDVCCCKPQELPPGQEGELLLGFFQKAVAPELL 841
DB 895 LSALESLTSPALAGL-----LLRPR---PAG---GAVAGRFOR-----LL 929
QY 842 HWITR-----GWVLLFLALFGVSLYSMSCHISVGLDQELALPKDS 881
DB 930 QVLGRPLNAPAEYGNVAKVVRVSGLAIVYVGLLGLTW-----VGFG---AVFP-- 977
QY 882 YLLDYFLFLNRYFEVG-----APVYFVTTLGVNFSSEAGNNAICSSAGCN 927
DB 978 ---GFVPMQDKYLVGIAQLPNGASLERTDAVVRQMSRIGLD---EPGVESVAVFPGLSV 1031
QY 928 SFSTOKIQYATEFPQSVLAIPASSWVDDFDWLTPSSCCRLYISGPNKDKFC----- 980
DB 1032 NGFVNVNPAAMVF-----FMDLFES-----RTSADLGAIAIAGRLQARFASPDG 1079
QY 981 ----PSTVNSLNLCKNCMSITMGSVRPSVEQPHKYLPMFLNDRPN-----I 1022
DB 1080 GVFPFPPVPGLG-----TIGGFMQVE-----DRGGAGLEALARQTQVLWM 1120
QY 1023 KCPKGG-LAAYSTSVNLT-----SDQVLSRFMAY----- 1052

Db 1121 KATESGQLGLMTSFDINAPQLEVVVDRTKVKSGQVRLADVPFALQVILGSLYINDPFRF 1180
Qy 1053 -----HK-----PLKNSQDYTEALRAARELAAN--ITA 1078
Db 1181 GRITYKVTQAQADAPHRQAEAGRLQVNRNAGAMPLFSSPVTVPSSGPRVHYNGPSA 1240
Qy 1079 DLR--KVPGTD-----AFVFP-----YTIITNVFYQVLTILPEGLFMLSICLVP 1122
Db 1241 DISGALPGVSSQAVAMERLAGEVLPEGMTPEWTDLTVOQKLA--GNSALFIFPLCVL- 1298
Qy 1123 TFAVSCLLGLDLRSGLN---LLSIVMILVDTVGPALWDISYNAVSLINLVSAGMS- 1178
Db 1299 ---LAYLILAAQYNSWLLPLAVLILVPMCLLSAIAAGVWLMGGDNNVFVOIGLVVLVGLAA 1355
Qy 1179 -----VEFVSHITRSFAISTKPTWLERAKEATISMSAVFAGVAMTNL---PGILVLGL 1229
Db 1356 KNAILLVEPARTLEAGA-----RALEAVEACRLRLRILMTSLAFTAGVVPFLVM 1406
Qy 1230 AXA-----QLIQIFPRLNLLTLGLLHGLVFLPVI 1261
Db 1407 ASGAGAEQMGVAVFAGMLGVTFLGLFTFPVYVLV 1444

RESULT 66

US-09-252-991A-16989
; Sequence 16989, Application US/09252991A
; Patent No. 6551795

GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 16989

; LENGTH: 1051

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-16989

Query Match 2.0%; Score 140.5; DB 2; Length 1051;

Best Local Similarity 20.5%; Pred. No. 0.0022;

Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

Qy 506 NOTLMGOT--SQVDMKDHFLYCANAPLTFKGTALALSCWADYGAPVFPFLAIGGYKGD 563
Db 236 NATIIGKTLQVDAQENILLKVN-----PDGSQVRLKDVADV-----LG---GQD 279
Qy 564 YSEAEALIMFTSLNNYPA-----GDPLRAQAKLWEE--AFLEEMRAFORRMAGMFQ 612
Db 280 YS-----INAQFNGSPASGAIKATGANALDTAKAIQTIANLEPFPQGMKVYPYD 333
Qy 613 VTFTAERSLEDEINTTAEADLPFATSYIVIFLYISLALGSYSSWRVMDSKATL--GL 670
Db 334 TTPVVSASTH-EVVKTLGE-----AILLVELVMYFLQNF-----RATLIPTI 375
Qy 671 GGVAVVLG--AYMAAGPFSYLGRSSLVILQVVPFLVLSVG--ADNIFIFVLEVRQLPR 726
Db 376 AVFVLLGTFTGVLAAPG-----SINTLTFMG-MVLAIGLLVDDAIVVVENVERVNA 426
Qy 727 RFG-EPREV---HIGRALGR-VAPSMLLCSLSEACFFLGLALTPMPAVKTFALTSLAVI 781
Db 427 BEGLSPREARAKSMQIQCALGVIAWLVSAVFLPAFFGSGTGV--YQFSITIVSAMA 484
Qy 782 LDFLLQMSAFVALLSDSKRQEARLDVCC--VKPQELPPPGQGBGLLLGFFQKAYAP 838
Db 485 L-----SVIVALI-----LTPALCATMLKPIEKGDHKGKGFPGFNRMFUS 527

Qy 839 FLHMITRGV-----LLLFLALFGVSLYSMCHISVGL-----DQ-----ELALPKD 880
Db 528 -TTHGVERGVASLTKHRAVYLLIYVIVAGMIWMFTRIPTAPLPDDEQGVLPFAVQTTPG 586
Qy 881 S-----YLDYFLFLNRYFVGAPVYFVTTLYGNFSS- 913
Db 587 SSAERTQVVVDMSRREYLLK-----ESSSVSSVFTVTGPFAGRGQSSGMAFIMLKP 638
Qy 914 -----AGMAI---CSSAGCNFSTQKIQYATEPPEQSVLAIPASSWDDDFD----- 959
Db 639 WEERPGGNSVPELAKRAQMHFFSPKDMVFA--FAPPSVLELGNATGDFLFLQDQAGVG 696
Qy 960 ----MLTPSSCCRLYISGPNKDKFCPSIVN-----SLNCLKNCMSI 996
Db 697 HEVLLQARNKFLMLAAQNALQVRPNCHMSDEPOYKLEIDDEKASALGVSLADINSTVSI 756
Qy 997 TWGSV-----RPSVEQPHKYLPMFLNDRPN-----IKCPKGGLA---AYST 1034
Db 757 AMGSSYVNDPIDRGRVVRV-----YLOGRPDARMNPDDLKMWYVRNDKGMVFPFNAPAT 810
Qy 1035 -----SVNLTSDGQVILASRFMAYHKPLKNSQDYTEALRAARELAANITADLAKVPGTDP 1088
Db 811 GKWEYGPXKLYRNGVPAIMEILGEPAPGLSSGD---AMAAVEE---IVKQLPKGVG--- 860
Qy 1089 AFEVFPYITITNVFYEQYLT-----ILPEGLFMLSIC-----LVPTFAVS 1127
Db 861 -----YSWTGLSYERLSGSOAPALYALSILVVFLCLAAALYESWISIPFSVMLVPLGVI 914
Qy 1128 CILLGLDLRSLNLLSIVMILVDTVGFMALWDISYNAVSLINL-----V 1172
Db 915 GALLATSMR-GLSNDVFFQVGLLTIGLSA-----KNAILVEFAKELHEQKGVIEAAI 968
Qy 1173 SAVGMSVFEVSHITRSFAISTKPTWLERAKEATISMSAVFAGVAMTNLPGILVL-GLAK 1231
Db 969 EACRMRLRPVMTSLAFTILGVVP-----LAISTG-----AGSGSQAHTGTGIGMVT 1016
Qy 1232 AOLIQIFPRL 1242
Db 1017 ATVLAIFWPL 1027

RESULT 67

US-09-328-352-4877

; Sequence 4877, Application US/09328352

; Patent No. 6562958

; GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER

; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC99-03PA

; CURRENT APPLICATION NUMBER: US/09/328,352

; CURRENT FILING DATE: 1999-06-04

; NUMBER OF SEQ ID NOS: 8252

; SEQ ID NO 4877

; LENGTH: 1005

; TYPE: PRT

; ORGANISM: Acinetobacter baumannii

US-09-328-352-4877

Query Match 2.0%; Score 140; DB 2; Length 1005;

Best Local Similarity 20.0%; Pred. No. 0.0023;

Matches 167; Conservative 128; Mismatches 328; Indels 212; Gaps 40;

Qy 569 ALIMFTSLNNYPAGDPLRAQAKLWEEAFLEEMRAFORRMAGMFQVTTFTAERSLE----- 622
Db 105 ALNPTY-NNSVVSQDQTA-AILL-LEESPEGFQKMMGPINKIVSESEQKDMTISVG 159
Qy 623 -DEINRTTAEDLP-----IPATSYIVIFLYISLALGSYSSWRVMDSKATLGLGVAVV 676
Db 160 GNPVYLDKAEADYSKRINILFPPIAVLVLGL---LHFAFRSKQGLILPLVTAL-----LAVA 212
Qy 677 LGAVMAANGFTSYLGRSSLVILQVVPFLVLSVGADNIFIFVLEY-----QRLPRRPGEP 732

Db 213 WG--MGWGLFK-----QPMDFNSFTPIILIAAGHAVQLKRYEYDFDRLIAQMEPK 266
QY 733 EVH-----IGRALGRVAPSMMLCSLSBAICFFIGALTPMPAVRTFALTSGLAVILDFLOMS 789
Db 267 AANSEAVQSLVRVGVFVMMVLAGGIAAGFFSLLTFTNIPTIRSFSGIFTGIGISTIVIENT 326
QY 790 AFVALLSLSKQERASRLDVCCVKPQBLPPG-----QEGLLLGFFQKAVAPLLHWIT 845
Db 327 FIPALRSM-----LPPSVVKVRKGL-----P-IWDWIP 355
QY 846 RGV-----VLLFLFALFGVSL--YSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893
Db 356 NRIGDIVLSVRPMLMTAIAAMGILLAIAGTSIRIVVNDNRNFFSBDLPQQDDRFNQ- 414
QY 894 FEVGAPVYVVTTLGYNFSEAGMNAICSSAGCNSFTQKIQYATEFFEPEQSYLAIPASSW 953
Db 415 -----SLGCTNSL-----YIMVDTKVRDGIENPE----- 438
QY 954 VDDFDLWLTSPSCCRLYISGPNKDKFCS-----TVNSLNLCKNCMSITWGSVRPSVEQ 1007
Db 439 ILKALD-----NTEKPAHSIPEVGKTIISIVDIYKR-MNQAMNADQPAFQ 482
QY 1008 FHKYLPWFNLDRNPKCPKGGIAAYSTSVNLTSDQVLAS-----RFMAYHKPLKNSQD 1061
Db 483 -----VPATKDVVAQYLLLYSMSGEPTDFDSYIDTTQRYAKITILLKTGSN 528
QY 1062 Y--TEALRAARELANITADLRKVPCTDPAFVFPYTTNNVFEQYVLTLPBGLFMLSCL 1119
Db 529 HRIKEILSLKTYMAGQLGDKAVVSGFDVTQTIALTETW-HGKLMNL-----QISFA 582
QY 1120 LVPTAVSCLLGLDLRSLGILLNLSIVMLTDTVGFMALWDSYNVAVSLNLSVAGMSV 1179
Db 583 V---PFISALVF-RSISAGLVLTLPLFSILAIFGVMGLDIPNLINSLISAMAVGIGA 638
QY 1180 EF-----VSHITSPALSTKPT---WLERAKEATISMGSAVFAGVAMTNLPGLVL-- 1227
Db 639 DYAIYFLYRLREILREEGDIDKAIKTLSTAGKASLFVATAVAGY-----GVLSLQ 692
QY 1228 GLAKAQLIIFFRNLTLTLGLHLGVLPVLVSYGPDVNPALALEKRAEBAVAV 1287
Db 693 GFHVHQLAMFI-----VIAMFSVFATLIMVPTMLI-----LKPRFIFSSKKKSIPVAQT 744
QY 1288 MWAS-----CPNHSRVSADNI--YVNHSEGS-IKAGAGISNPL--PNNGRQ 1331
Db 745 VVTSLLLTALTMSMPTKSHADEVDIVNRSDDAKFSLSTASAKFILTSKNGEQ 799

RESULT 68
US-09-902-540-12490
; Sequence 12490, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12490
; LENGTH: 826
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-12490

Query Match 2.0%; Score 138; DB 2; Length 826;
Best Local Similarity 19.4%; Pred. No. 0.0025;
Matches 134; Conservative 116; Mismatches 253; Indels 187; Gaps 33;

QY 613 VTTEASRLEDEINRTTAEDLPIFATSVIVIFL-YISLALGSYSWSRVMVDSKATILGLG 671
Db 242 VTGTFKKIDQ--RVIVGDLASASTLAWLLLAYLAFHPSALS----- 284
QY 672 GVAVLGAVMAAGF-FSYLGIRSSVLQVVPFLVLSVGADNI--FIFVLEYQRLPRRP 728
Db 285 -VGLTMAPVMAGLQWTFGVGLAYQVNL-LTGFLAAVLGGLGVEHGIHLGRYTTLRSE 342
QY 729 GEPREVIHGRALGRVAPSMMLCSLSSEALCPFLGALTMPAVRTALTSGLAVILDFLOM 788
Db 343 GMSSEAVDESFRHTGSSLLAALVAALTFLSLAWSBFRAPREFGIIAAGMLVSIYSV 402
QY 789 SAFVALLSLSKQERASRLDVCCVKPQBLPPGQEGLLLGFFQKAYAP--FLIHWITR 846
Db 403 LLLPALLGLATRFQMSRV-----QEG-----AAGPLSLLARWLPR 439
QY 847 GVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYVVTTL 906
Db 440 -----SYRGVGIV---VGVG---LALVSQAY-----RI 462
QY 907 GYNFSEAGMNAICSSAGCNSFTQKIQYATEFFEPEQSYLAIPASSWVD----- 955
Db 463 SFNYSRTLEDYKQASA-----VLDQKNDILGYSQTPVVVLTDSQEMEREVVRQLEARK 517
QY 956 -----DFI---DMLTPS-----SCCRLYISGPNKDKFPCPSTVNSLNLCKNCM 994
Db 518 VARGKESTIDFVGALDVLVPRQDEKQAILQASAKLEKLDPERLPEDTRNL-----VR 572
QY 995 SITMGSVRPSVEQFHKYLPWFNLDR-PNICKPKGGIA-AYSTSVNLTSDGQVLASRFMAY 1052
Db 573 ALNMAKAPFTQE---ALPTSVRHQFESLDGSTGGVVLVYAGGVS-LSDGE--GTRKFA- 625
QY 1053 HKPLKNSQ--DYTEALRAARELANITADLRKVPCTDPAFVFPYTTNNVFEQYVLTLP 1110
Db 626 -KEVRGLQMPGQSQVSAAGEAL--ILADIIDM-----VSR 657
QY 1111 EGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISIVMLTDTVGFMALWDSYNVAVSLIN 1170
Db 658 EGPRILAAVLSVLVAMWLTIG-KLRTALICWLPILLSVGLVGLMSLLGLQFNVLNIMV 716
QY 1171 LVSAGMSVFEVSHITRSFAISTKPTWLERAKE---ATISMGSAVFAGVAMTNLPGLVL 1227
Db 717 LPVLVGTVDAGVHLVQRLG-----ERGADFVSVAETGRAIMGGL-LTSAIGFVAL 767
QY 1228 GLAK-----AQLIQIFFRNLTLTLGL 1250
Db 768 ILAKHGLNSIGDLANT--GFGINMVIVLGL 796

RESULT 69
US-09-489-039A-10700
; Sequence 10700, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10700
; LENGTH: 1040
; TYPE: PRF
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10700

Query Match 2.0%; Score 137.5; DB 2; Length 1040;
Best Local Similarity 18.1%; Pred. No. 0.0042;
Matches 162; Conservative 127; Mismatches 289; Indels 315; Gaps 41;

Db 604 NIDKAVELYSSGERKLANVFDLETMF-REYNTYAQISGLRKPDPNSQRNDRRIIQDF 662
QY 630 AEDLPATSYIVIFLYISLALGSSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFFSY 689
Db 663 SEILADLGSIGKVI---VNVASGAPSLFGGIVT-----GILNFIKPNLGLGMFTFL 710
QY 690 LGIRSSILVILQV 701
Db 711 IGAIVILVILV 722

RESULT 74
US-08-720-229-14
; Sequence 14, Application US/08720229
; Patent No. 6022542
; GENERAL INFORMATION:
; APPLICANT: Rose, Timothy M.
; APPLICANT: Bosch, Marnix L.
; APPLICANT: Strand, Kurt
; TITLE OF INVENTION: GLYCOPROTEIN B OF THE RFHV/KSHV
; TITLE OF INVENTION: SUBFAMILY OF HERPES VIRUSES
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/720,229
; FILING DATE: 26-SEP-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Schiff, J. Michael
; REGISTRATION NUMBER: 40,253
; REFERENCE/DOCKET NUMBER: 29938-20002.00
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 808 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-720-229-14

Query Match 1.9%; Score 130.5; DB 2; Length 808;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 92; Conservative 71; Mismatches 162; Indels 107; Gaps 18;

QY 323 TSLSKLS--FSTHLLG--QFQGGTWWASWPIITLVLSVPPVVALAAGLVFTLETDD 378
Db 345 TTLEDKIRAVNTHTPNGTAYYYQTEGGMILVQ-----PLIAIELEAMLEATTS 395
QY 379 PYVELSAPNSQARSEKAFHQHGFPRFNQVILTPAPRSSYRYSLLGPNFSGILD 438
Db 396 PVP--SAPTSRSKRRAIRSDVSAGSENNVFL---SQIYAYDKLROSINN----- 444
QY 439 DLILLELLEL---QERLRHQVMSPEAQRN-LSLQDICYA-PLN-----PDNTSLYDCINS 489
Db 445 --VLEELAITWCBEQVROTWWYELAKINPTSVMTAIYKPVSRKALGDVISVTEC---- 498
QY 490 LLOYFQNNRTLLLTANQTMQGTQVQWDKHFLYCANAPLTFKO----- 534

Db 499 -----INVDOSSVSIHKSLKTENNDCYSRPVPVTFKFVNSSQLFKGQLGARN 545
QY 535 ----GTALALSCWADYGPVFPFLAIGGYKGKDYSEAE-----ALIMTF-- 574
Db 546 ELLLESUVENC--HQNAETFTAKNETHYFNKYHVHETLPVNNISTLTDTFLALNITFIE 603
QY 575 -----SLNNYPAGDPRLAQAKLWBSAFLEEMARFORMAGMFQVTFATERSLEDEINRTT 629
Db 604 NIDKAVELYSSGERKLANVFDLETMF-REYNTYAQISGLRKPDPNSQRNDRRIIQDF 662
QY 630 AEDLPATSYIVIFLYISLALGSSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFFSY 689
Db 663 SEILADLGSIGKVI---VNVASGAPSLFGGIVT-----GILNFIKPNLGLGMFTFL 710
QY 690 LGIRSSILVILQV 701
Db 711 IGAIVILVILV 722

RESULT 75
US-09-252-991A-24533
; Sequence 24533, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24533
; LENGTH: 1069
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24533

Query Match 1.9%; Score 130.5; DB 2; Length 1069;
Best Local Similarity 21.2%; Pred. No. 0.021;
Matches 123; Conservative 80; Mismatches 208; Indels 169; Gaps 27;

QY 815 PQELPPPGOGEGLLGLLPQKAVAP--FLHWTTRGVVLLFLALFGVSLYSMCHISVGLD 872
Db 19 PGAAPEPGSGGG-----AAVTPRAGISGWCIRHPITATLITLASLLGLLAFRLGV- 71
QY 873 QELALPKDSYLLDYFLNRYPEVGAPVVF---VTTLGYNFSSBAGMNAICSSAGCN-- 926
Db 72 --APLPEADPPT---IQINALLPGSGPETMASSVATPLEVQSAIPGITEMTSSSALGTT 126
QY 927 ----NFSFTQKIQYATPEFQSYLAIPASSWWDDFDMLTPSSCCRLYISGN----- 975
Db 127 TLTLQFLSKSIDVAA--QEVQAAINAAAG-----RLPVDMENTLPTWRKV 169
QY 976 -----KDXECP-----STVNSLNLCKNCSITMCSVSPSVSEVQ 1007
Db 170 NPADSPIMILRVNSEMPLIELSDYAETILARQLSQVNGVQI-----FVVGQORPAI-- 222
QY 1008 PHKLPWFNDRPNKCPKGGLAAY-----STSVNLT-----SDGVQLASRFM 1050
Db 223 -----RLOAQPEKLAAYQLTLADLRQLQSASVNLAKALYGEGRVST---L 266
QY 1051 AYHKPKNSQDYTEALRAARELANITADLRK-----VPG----- 1085
Db 267 AANDQLFNASDYDDLWAVRQCAPVFLKDVARIVSAPDDYVQAWPNGVPGVALVILRQP 326
QY 1086 ----TDPA-----FEVPPYTI-TNVPFEQVLTILPEGLFMLSCLVTFPAVSCLL 1130
Db 327 GANIVDTADAIOAALPRLREMLPATIEVDVLNDRRTI-RSSLHEVELTLLLTIGLVILV 385

; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2020
; LENGTH: 801
; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: amino acid sequence

US-09-710-279-2020

Query Match 1.8%; Score 126; DB 2; Length 801;
Best Local Similarity 19.9%; Pred. No. 0.033;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

```
QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 62 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 121
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTLGYNFSEAGM 916
Db 122 MGVVLSDFNFIILYLFWEILTSSPFLISFWREKASIIYCAQKSLIITVLG-GLSMLGGI 180
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 181 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVMIEMIGATKSAQVFFYIWLDPAMEA 238
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 239 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATKQ 294
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAVHKPKNSQDYTEAL 1066
Db 295 HDLK-----GILAFSTVSQLGTMISMLGIGAVSYHYQGANSQLYVAGF 337
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPFYTITNVFYEQVILT 1107
Db 338 VAAIPLHNLHATFKGALFMTGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 387
QY 1108 ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 388 VIT-----TLSMAGVPPP--NGFLSKEKFLSINVTNHLNLSINTLG-----ILLPIIA 435
QY 1168 LNLVSAGVMSVEFVSHITRSFAISTKPTWL-ERAKEATISMGSAVFAGVAMTNLPGLIV 1226
Db 436 IIGSITFTVYSIKFILHI---PFGSYKPEALPQAHSSILM---LISPIILTSL--VIV 487
QY 1227 LGLAKAQIQTIPFFRLNLLITLLG-----LLHGLVFLPVLISYVG 1266
Db 488 FGLFPFSILTQSIIIEPASVAVSQTSNITAEFHLPHGIT--PAFLSTIG 532
```

RESULT 78

US-09-134-001C-3105
; Sequence 3105, Application US/09134001C
; Patent No. 6380370

; GENERAL INFORMATION:

; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C

; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964

; PRIOR FILING DATE: 1997-11-08

; PRIOR APPLICATION NUMBER: US 60/055,779

; PRIOR FILING DATE: 1997-08-14

; NUMBER OF SEQ ID NOS: 5674

; SEQ ID NO 3105

; LENGTH: 808

; TYPE: PRT

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3105

Query Match 1.8%; Score 126; DB 2; Length 808;
Best Local Similarity 19.9%; Pred. No. 0.034;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;

```
QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 69 LNMHPHIGMNFNLYVDGLGLFLSLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 128
QY 874 ELALP-KDSYLLDY-----FLFLNRYFEVGAPVY-----FVTLGYNFSEAGM 916
Db 129 MGVVLSDFNFIILYLFWEILTSSPFLISFWREKASIIYCAQKSLIITVLG-GLSMLGGI 187
QY 917 NAICSSAGCNNSFTQKIQYATEPPEQSYL-----AIPASSWVDDFDW 960
Db 188 --ILLSLATDTFSIQAMISKASDIQNSPPFILLVMIEMIGATKSAQVFFYIWLDPAMEA 245
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNLKNCMSITWGSVRPSVE 1006
Db 246 PTPVSAYLHSATMVKAGLYLIARITPFAISEGWWTIT-----LVGLITLFWASLNATKQ 301
QY 1007 QPHKYLWFLNDRPNIKCPKGLAAVSTSVNLTSDQVLASRFMAVHKPKNSQDYTEAL 1066
Db 302 HDLK-----GILAFSTVSQLGTMISMLGIGAVSYHYQGANSQLYVAGF 344
QY 1067 RAARELAANI-----TADLRKVPGTDPAPFVFPFYTITNVFYEQVILT 1107
Db 345 VAAIPLHNLHATFKGALFMTGGIDHSTGTRDVKKLGG--LLTIMPISFT-----LT 394
QY 1108 ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSVIMLVDTVGFMAWDISYNAVS 1167
Db 395 VIT-----TLSMAGVPPP--NGFLSKEKFLSINVTNHLNLSINTLG-----ILLPIIA 442
QY 1168 LNLVSAGVMSVEFVSHITRSFAISTKPTWL-ERAKEATISMGSAVFAGVAMTNLPGLIV 1226
Db 443 IIGSITFTVYSIKFILHI---PFGSYKPEALPQAHSSILM---LISPIILTSL--VIV 494
QY 1227 LGLAKAQIQTIPFFRLNLLITLLG-----LLHGLVFLPVLISYVG 1266
Db 495 FGLFPFSILTQSIIIEPASVAVSQTSNITAEFHLPHGIT--PAFLSTIG 539
```

RESULT 79

US-09-252-991A-27976
; Sequence 27976, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27976

; LENGTH: 1123

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27976

Query Match 1.8%; Score 125.5; DB 2; Length 1123;
Best Local Similarity 22.1%; Pred. No. 0.067;
Matches 108; Conservative 82; Mismatches 191; Indels 107; Gaps 25;

```
QY 409 QVILTAPNRRSYRYDSLLGPKNFGILDLLLELLELQERLRLHQLVMSPEAQRNISIQ 468
Db 141 QLIVLTLS--SSGRYDAMDGLGIASSNV-----LQALRRVEGVGVKVTWGAETVRIWPD 193
```

QY 469 DICVAPLNPDNTSLYDCINSLLQYFQNNRTLL-----LLTANQTLMGQTSQVD-----WK 519
Db 194 P---AKLTSMLSASD-LVNVRHR--NARLTGVDIGNLGVDPDSAPISATVKVDDTLVTP 247
QY 520 DHFLYCANAPTFK-DGTALALSMA--DYGAPVPFLAIGYKGYKGYSEBALIMTFSL 576
Db 248 EQF---GEIPLRADGGAIRLRDVARVERFGQSEYGFV-----SRVQMTATGLA 294
QY 577 NNPAGDPRLAQAKLEBAFLEEMRAFORMAGMFQVTTFAERSLEDBINRTTAEIDLPIF 636
Db 295 VKMAPGSNAVATAKRIRATLDELSRYFPEGVS--YNIPTYSAFVEISIRKVVSTLLEAM 352
QY 637 ATSYVIVFLYLSLALGSSYSRVMVDSKATILGLGGVAVVLGAVMAAGPFS-YLGISS 695
Db 353 LLVFAVMYLF-----NONFRATL-----IPTLVVPVALLGTFTVMVLGLGFS 393
QY 696 LVILQVVPFLVLSVG--ADNIFIFVLEYORLPRPG-EPREVIHGRALGRVAP-----SM 747
Db 394 INVLMFG-MVLAIGILVDDAIIVVENVERLMAEEGLSPHDATV-KAMRQISGAIVGITV 451
QY 748 LICSLSAICFIFLGTALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKROEASRL 807
Db 452 VLVSFVFPMAFSGAVGNI--YRQPAVT--LAVSIGF---SAFLAL-----SLT 493
QY 808 DVCCCKVQBELPPPGQEGELLGFTQKAVP-----FLLHWITRGV 848
Db 494 PALCATLRLPIDADHHEKRGFGFENRPAFLRLTGRYRNAVAGILARPIRWMLVYALVIGV 553
QY 849 VLLFLAL 856
Db 554 VALLFVRL 561

RESULT 80

US-09-252-991A-28767
; Sequence 28767, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28767
; LENGTH: 789
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28767

Query Match 1.8%; Score 123.5; DB 2; Length 789;
Best Local Similarity 21.9%; Pred. No. 0.056;
Matches 154; Conservative 98; Mismatches 240; Indels 211; Gaps 41;

QY 635 IFATSVIVIFLYISLALGSSYSRVMVDSKATILGLGGVAVVLGAVMAAGPFSYLGIRS 694
Db 236 VMVVGFGGCLLITLVL---LYWFTKCI--RSTIAV--LITTLVAVLMQLGULLNVGFLG 288
QY 695 SLVILQVFPFLVLSV-----GADNIFIFVLEYQRLPRP-----PGEPRE 733
Db 289 DPYSM-LVDFLFIAGISGHQVQKINGIALQSSGADN-----ALMAARLTFRQLFLPG-----339
QY 734 VHIGRALGRVAPSMILCSISEAICFFLGTALTPMPAVRTFAL--TSGLAVIDL--DFLLQMSA 790
Db 340 -----MTAILADAVGFTITLLVIDIGVIRELAIGASIGAVIVFTNLLIPV 385
QY 791 FVALLSLDSKROEASRLDVCCCKVQELPPPGQEGELLGFTQKAVAPFLLHWITRGVVL 850

Db 386 AISVIGISKAVQSKDDAV-----REHP-----FWRLLSNFASPKVAP-----VSI 427
QY 851 LLFLALPGVSLYSWCHISVG-LDQ---ELALPKDSYLLDY-FLFLNRYFVEGAPVYFVTT 905
Db 428 AIALMLLGGGLWYKGLKIGLDQCAPELR-PDSRYNLNDNDFIIRN-----472
QY 906 LGYNFSSSEAGNVAI-CSSAGCNNSFTQKIQYATEFPPEQSYLAIPASSW-----VDDF 957
Db 473 --YSTSSDVLVVMVKTSPGCGST-----HQAAMDELAWKLENTGVSQA 516
QY 958 IDMLTPSSCCRLYISGPNKDKFCPSTVN-SLNCCLKNCSITWG-----SVRPSVEQPHK 1010
Db 517 ISMTVVS---KQVIGKWNENGLKWETLSRNDQVLNNSIARADGLYNTDCSLAP-----566
QY 1011 YLPWFLNDRPNKCPKGGGLAAYSTSVNTSGOVLASRFMAYHKPLKNSQDYTEALRAAR 1070
Db 567 -LLVFLNDHK-----AETLDRAVAAVQAFAAENDKDLQFOLAAGNAGI 609
QY 1071 ELAANITADLRKVPCTDPAFEVFPYTTITNVFEQYLTILPGLFMLSLCLVPTP---AV 1126
Db 610 EAATN-----EVIKQSELIILVLVVICV-----AAMCMI--TFRSPAAT 646
QY 1127 SCLLGLDLRGLLNLLISIVMLVDTVGFMAWDISYNAVSLINILVSAGMSVEFVSHI- 1185
Db 647 LCIVLPLILTSVGNAL-----MAALGIGVKVATLPVIALGVGIGVGYIYI 694
QY 1186 TR--SFAISTKPTWMLERAKEATI-SMGSAV-FAGVAMTNLPGILVLGLAKAQLIOI-PFF 1240
Db 695 TRLESFLRMGLP--LQEAYYETLSTGKAVLTGLC-----LAIGVATWTFSAIKFOA 745
QY 1241 RLNLIT---LLGLLHGLVFLPVILSYGVPDVPNPALALEOKRA 1280
Db 746 DMGLMLTFLMLNMFGALWLLPALARFL---INPA-KLOARKA 784

RESULT 81

US-09-328-352-8064
; Sequence 8064, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 8064
; LENGTH: 1065
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-8064

Query Match 1.8%; Score 123.5; DB 2; Length 1065;
Best Local Similarity 17.5%; Pred. No. 0.094;
Matches 200; Conservative 178; Mismatches 409; Indels 359; Gaps 54;

QY 344 WGTW-VASWPLTILV-LSVIPV-----ALAAGLVFTTELTTDPVELMSAPNSQAR 391
Db 43 WGTWAVKNTFVDALPDLSDVQVIIRTNFSGQAPQIVENQVITYPLTT---TMLSVPGVKTV 99
QY 392 SEKAP-----HDQHPGPFRTNOVI-----LTAPNRS-----XYDS 424
Db 100 RGYSPFGDSFVVYIFDEHTDLYWARSVLEYLNQIGKMPANAKSSLGDPATGVGVYIEY 159
QY 425 LLLGPKNFGILDLDLLELLELQERHLQVWSPQAORNI SLQDI CYAPLNPNDNTSLYD 484
Db 160 ALVDP---TGQHDL-----SELRSIQDFWFLKYELK-TLPNVAEVATIGGMVKQYQ 205
QY 485 CCIN-----SLLOYFQNNRTLLLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGKTALAL 540
Db 206 VULDPSKMATLGVTQNNVIEAIQKANQETGGSVLEMAETEYVMVRASGYLKTLDLDFRQIPL 265

```
QY 541 SCHADYGAPVFPPLATGYGKDYSEBALIMTFSLNNTYPAGDPRLQAQKWEAEPLFM 600
Db 266 RTNS-FGVFV-:::--VATIQG-----PEMRG-----ISBL 295
QY 601 RAFOREMAGWF-----QVTFABRSLEDEINRTT---AEDLPFATSYIV-----642
Db 296 NQGETVGGVILRAGKNAREITAVAKLAHQSLPKGVQVVPYDRLSDRAVENL 355
QY 643 -----IFYIYSLALGYSYSSWRVMDSKATLGLGGAVALVAVMAAMGFFSYLGIRSSL 696
Db 356 SHKLIEEFIVLVCGLFLWHLRSAMVAIVSLPLG-----ILSAFLVMHYQGLNANI 407
QY 697 VILQVVPFLVLSVGA-----DNIFIFVLEYQRLPRRPGEPREHV-----IGRALGR 742
Db 408 MSLGGI---AIAIGAMVMDAAIWMVENAHKHIEITWQH--BHPQVLEAQVRWNIITRSASE 462
QY 743 VAPSMILC---SLSAICFFLGA-----LTPMPAVRTEAL--TSLAVILDFLOMSAF 791
Db 463 VQPTLFFCULLIITLSPITFOAGRGRLFSPLAFTKTYAMAAAGLSITLIPVL-MGYM 521
QY 792 VALLSLDSKQBSRLDVCCVKPQELPPPGQEGLLGFFOKAYAPFLHWHITRGVILL 851
Db 522 I-:::--RGKL-----PSQRNPLNRFELIKIYSPMLDKVLAHPKTI 557
QY 852 L-FLALFGVLSYMSCHIS--VGLDQ-----ELALP-----KOSYLLDYFLNRYPE 895
Db 558 LGALLIFLISLPLFLRLGCEFLPNMDGDLLYMPSALPGLSAKASELIQ-----Q 608
QY 896 VCAPVYFVTTLGVNFSSEAGMAICSSAGCNPFSFKIQYATEPEQSYLAIPASSWD 955
Db 609 TDRMKTVEPATVTVGKAGRASATSDAPLEPFET--IOFK-----PRSEWRS 655
QY 956 DFIDMLTSSCCRLYISGNKDK-----FCPSTVNSLNCNKMSITMG-----999
Db 656 G---WTPDKLK-----ELDKAVQVPGLTNIWVPPINRDMLATGKSPIGIKVSAN 705
QY 1000 -----SVRPSVEQFHKYL-:::--WFLNDR-----1019
Db 706 DIQDIDRVAQIEQVAKQMPGVSSALAERLTGGRYVDVDINRMQAARYGLNIKDQVQIVS 765
QY 1020 -----PNIKPKGLAAVSTSYN-----LTSDQVLSRPMAYHK---1054
Db 766 SAIGENGETVEGLARYPINRYPREIRDSLEALRNLPILTSGQIVLSSVANIQTID 825
QY 1055 -----YTEALRAARELANITADLRKVGTPDPAFEVFPYTTITNVFYEQY 1105
Db 826 GPMLKSENARPSGHVYDV--QGRDL-ASVVQDLKQIDQ-----KVRSSAMSISYSGQ 878
QY 1106 LTILPEGLFMLSCLVPTFAVSCLLGL---DLRSGLNMLASIVMLVDTVGFMAIWDIS 1162
Db 879 PEFMERANARKVVIPITLMIIFLLYLIIFRQVQDAVLIMATLPALIGGIWMYLSDYH 938
QY 1163 YNAVSLINLVSAGMSVBF-----VSHITRSPAISTKPTWLRABKATISMSGA 1211
Db 939 FSAIVAGFIALAGAAEFVGMFLYKQAEHAQOQSLSSST-----ASLTEQQLNEA 991
QY 1212 VPAGVAMTNLPGILVLGAKAQLOIIPFRNLMLIT-----LLGLLHGLVFLPVI 1261
Db 992 IHTGAVLRVRPAMTVAVILAGLIPI-----LIGTGTGSELSMRIALPVMVGMSAPILL 1045
QY 1262 LSYVGF 1267
Db 1046 SMFVIP 1051
```

RESULT 82

```
US-09-252-991A-21000
; Sequence 21000, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
```

```
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21000
; LENGTH: 1033
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21000
```

Query Match 1.8%; Score 123; DB 2; Length 1033;

Best Local Similarity 26.2%; Pred. No. 0.1; Matches 70; Conservative 39; Mismatches 110; Indels 48; Gaps 11;

```
QY 1017 NDRPNKCPKCGGLAAYSTSVNLTSQVLAASFMAHYKPKNSQDVTEALRAARELANI 1076
Db 276 NDPPAPRMRFWGDEAIGLAVAMKPGDIL-----VLKALE-----TEPARLQOQSLPAGL 325
QY 1077 TADLRKVGTDPAPFVFPYTTINVFYEQVLTILPGLFMLSCLVPTFAVSCLLGLDUR 1136
Db 326 --ELKVSQDPAVR-----TGV--GBFIRVLAELAVIVLL-----VSPFSLG--LR 366
QY 1137 SGLNLLSIVMLVDTVGFMAIWDISYNVSLINLVSAGMSVBFVSHITRSPAISTKPT 1196
Db 367 TGLVVALSIPVLAMTFAAMHYFGILGHKISIGALVLAIGLLVDDAIITAVEMVAVMEQG 426
QY 1197 WLERAKETISMGSAVFAGVAMTNLPGILVLGLAKAQ-----LIQIFFRNLMLIT 1247
Db 427 Y-DRLKAAFAWTSTAFPMLTCTLTAAAGFLPINTAQSGTGEYTSLSFQVVTIAL-----480
QY 1248 LGLLHGLVFLPVLSYVGPVNPALA 1274
Db 481 VVSWFAAVFVFP-----YLGAKLPLDLA 503
```

RESULT 83

```
US-09-489-039A-13014
; Sequence 13014, Application US/09489039A
; Patent No. 6610836
```

GENERAL INFORMATION:

APPLICANT: Gary Breton et. al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 2709.2004001

CURRENT APPLICATION NUMBER: US/09/489,039A

CURRENT FILING DATE: 2000-01-27

PRIOR APPLICATION NUMBER: US 60/117,747

PRIOR FILING DATE: 1999-01-29

NUMBER OF SEQ ID NOS: 14342

SEQ ID NO 13014

LENGTH: 1049

TYPE: PRT

ORGANISM: Klebsiella pneumoniae

US-09-489-039A-13014

Query Match 1.8%; Score 123; DB 2; Length 1049;

Best Local Similarity 18.2%; Pred. No. 0.1; Matches 146; Conservative 107; Mismatches 244; Indels 304; Gaps 34;

```
QY 624 EINTTAEPLPFPATSYIVIFLYISLALGYSYSSWRVMDSKATL--GLGGVAVVLG--A 679
Db 340 EAVKTLVE-----AIIIVELVMYFLQNF-----RATLIPTIAPVVLGTF 382
QY 680 VMAAGPFYSLGIRSSLVILQVVPFLVLSVG--ADNIFIVLEYQRLPRRPGEPREHVIG 737
Db 383 VLAAGFG-----SINTLTMFG--WLAIGLLVDDAIIVVVENVERVMAEEGLPPKBEATR 433
QY 738 RALGR-----VAPSMILCSLSEACFFLGAITPMPAVRTALTSLAVILDFLLQMSAFV 792
```

Db 434 KSMQIQALVGIANVLSAVFIPMAFGSGTGAI--YRQSFITVSAMALSVLVALITLP 491
Qy 793 ALLSLDSKROBASRLDVCCCVKPOELPPPGGEGLLGFFQKAYAPFLHMI----- 844
Db 492 ALCA-----TMLKPIQKSGHGATTG-FFGWFNRMPDKSTHHTYDSVGNILR 536
Qy 845 TRGVULLFLALF-----GVSLSMCHISVGLDOE----- 874
Db 537 STGRYLVLIIIVGMAWLFVRLPSSFLPDEDOGVFL-SMAQLPAGATQERTQKVLDEMT 595
Qy 875 ---LALPKDS---YLLDYFLFLNR----- 892
Db 596 NYLTKEKNDVESVAFVNGFAGRGQNTGIAFVSLKDWSDRPGREENKVEAITARAMGYF 655
Qy 893 -----YFEVGAPVYFVTTLGVNFSSEAGMNAICSSAGCNFSFTQ----- 932
Db 656 SQIKDAMVFAFNLPAIVELGTATGDPDE-----LIDQGLGHEKLTQARNQLFGMVAQ 708
Qy 933 -----KIQYATEPPEQSYLAIPAS-----SWYDDFDLWLTLP 963
Db 709 HPDLVTGVRPNGLETPQFKIDIDQEKQAQALGVSDINTTLGAAMGSGSYVNDFFID---R 765
Qy 964 SSCCRLYISGPNKDFCPSVTNSLNCNKMSITMGSVRPSVEQPHKYLPLFLNDRPNIK 1023
Db 766 GRVKVYIMSEAKYRMLPEDIGK-----WYVRGSDGQM 798
Qy 1024 CPKGLAAYST-----SVNITSQGVLASRFMAYHKPLKNSQDYTEALRAARELANIT 1077
Db 799 VP---FSAFSTSRWEYSGPRLERYNGLPSLEILGQAAPGKST---GEAMALMEELAGKLP 852
Qy 1078 ADLRKVGPTDPAFEVPPYTIITNVFEQYLT-----ILPEGLFMLSCL----- 1120
Db 853 SGI-----GYDWTGMSYQERLSQGNQAPALYAI SLIVVFLCLALAYESWSIPP 899
Qy 1121 ----VPTFAVSCLLGLDLRSLNLLSVIMLVDTVTGFMALMDSYNVSLINL----- 1171
Db 900 SVMLVPLGVGVALAA--TFEGLTNDVYFQVGLLTIIGLSA-----KNAILIVEFAKDL 952
Qy 1172 -----VSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMT 1219
Db 953 MEKEKGLEATLEAVRMRLRPILMTSLAIFLGVMP-----LVISSG-----AGSGAQ 1000
Qy 1220 NLPGLVL-GLAKAQLIQIFF 1239
Db 1001 NAVGTGVMGGMVTATILAIF 1021

RESULT 84

US-09-252-991A-30591
; Sequence 30591, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30591
; LENGTH: 1052
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30591

Query Match 1.8%; Score 122.5; DB 2; Length 1052;
Best Local Similarity 17.6%; Pred. No. 0.11;
Matches 234; Conservative 166; Mismatches 400; Indels 531; Gaps 57;

Qy 134 SLFINVTR---VAQLGAGQLPAV-----VAYEAFYQHSFAEQSYSCSRVRVPAAATLAVG 186
Db 25 ALFISLAGLIVISKLPVQVNPVAPPQITITATYP-----GASAKVLVD 68
Qy 187 TMCQVYSGALCNAQRWLNFQO-DTNGGLAPLDITFHLLBPG----- 226
Db 69 SVTSVLESNGAKGLLYFESTNNSNGTAETVVTF---EPGTDPLAQVDVQNRLLKABE 125
Qy 227 ---QAV-GSGIQ-----PLNEGVARCNEOGDDVATCSCODCAASCAPAR 268
Db 126 RMPQAVLTQGLQVBQTSAGFLLIYALSKEGAQRSDTTALGDYAAARNINNELRLPGVGK 185
Qy 269 PQALDSFYLGQMPGSLVLIILCSVFAVTVILLVGRVAPARDKSMVDPKK-----GTS 324
Db 186 LQFFSSE-----AMRV-----WIDPQKLVFGLS 210
Qy 325 LSDKLSFSTHTLLGQFFQGMGTWVASWPLTILVLSVIPVVAL--AAGLVFTELTDDPVEL 382
Db 211 IDD---VSNAIRCONVQ-----VPAGAFGSAPGSSAQELTATLAVK 248
Qy 383 WSAFNSQARSEKATHDQHFQFFRTNQVILTAPNRSVRYDSLLGPKNFSGLDLDLL 442
Db 249 GTLDDPQBFQGVLRANQDGSVLRLADVARLELGEKESYNISRLNGTPTVGAQLSPGA 308
Qy 443 ELLS---LOBRLRLQVMSPEAQORNISLODI CYAPLNPNTSLY-----DCCINSLIQ- 492
Db 309 NAIQTATLVKORLAELSAPFE-----DMQYSV--PYDTSRFVDVAIEKVHTLIEA 358
Qy 493 -----YFQNNRTLL--LTANOTLMQOTSOVDKQHFLYCANAPLTFKQGTALAL 540
Db 359 MVLVFLGMFLPLQNVRYTLIPSIVPVCLLGLTLM-----YLLGFSVNMNTMFGMVLAI 413
Qy 541 SMADYGNAPVFPFLAIGYKQKDYSEAEALIMTSLNNYPAGDPRLAQAKLWEAFLEEM 600
Db 414 GILVD-----DAIVVENVERIMA-EGISPA-----EATVRAM 446
Qy 601 RAFQRRMAGMQVTFATERSLEDEINRTTABDLPI--FATSVIVIFLYISLAL----- 651
Db 447 KOVSGAIVGITLV-----LSAVFLPLAFWAGSVGVITYQQFSVSLAVSILFS 492
Qy 652 -----GSYSMSWR-----VMVDSKATLGLGGVAV 675
Db 493 GFALITFTPALCATLLKPIPEGHEKRGFFGAENRGFARVTERYSLLNSKLVARAGRFLM 552
Qy 676 VLGAVMAAGFFSYLGIIRSSLVILQVFPFLVSGADNIFIFVLEYQRLPRPGEPREVH 735
Db 553 VYAGLVAMLYGF-YLRLEPAFVPAEDLGYMWVDV-----QIP--PGASR--- 593
Qy 736 IGRALGRVAPSMLLCSLSEALCFILGALTMPAPVRTFA-----LTSGLAVILDFLQ 787
Db 594 -----VTDATGEELERFLKSREAVASVFLIS 620
Qy 788 MSAFVALLSLDSKROBASRLDVCCCVKPOELPPPGGEGLLGFFQKAYAPFLHMIWITG 847
Db 621 GFSF-----SQGDNAALAF-----PTFKDWSRG 645
Qy 848 VLLFLFLALFGVLSYSMCHISVGLDQELALPKDSYLL----- 884
Db 646 -----AEQSAABEIAALNEHFALPDGVTMAVSPPPINGLNGSGGFALRLMDR 693
Qy 885 -----DYFL-----FLNRYFE--VGAPVYFV-----TTLGYNFSSEAG 915
Db 694 TGVGREALLQARDTLIGEITQNPFLYAMMEGLAEAPQLRLIIDREKARALGVSFETISG 753
Qy 916 -MNAICSSAGCNFSFTQIKIYATFEPPEQSYLAIPASSWVDFFIDWLTPSSCCRLYISGP 974
Db 754 TLSAAGFSGEIVINDFTNAGRQQRVVIQAEQG-----NRMTPESVLELYV--- 796
Qy 975 NKDKFCPSVTNSLNCNKCMSITM--GSVRPSVEQPHKYLPLFLNDRPNIKCPKGLAAY 1032
Db 797 -----PNAAGNLVPLSAFVSVKWEQPV-----QLVRY-----NGYPSIRIVGDAAPGF 840

```
QY 1033 STSVNLTSDQVLSRPMAYHKLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEV 1092
DB 841 STG-----EAAEMERLAAQAPAGI-----860
QY 1093 FPYTITNVFEOYLTI-LPEGLFMLSICLVPTFAVSCLLGLDLRSLNLLSIVML--1149
DB 861 -GYEWTLGSLYOEKVSAGQATSLPALILV-----FLLVALYSWSIPUSMLIVPI 912
QY 1150 --VDTVGFMALWDISYNAVSLINLVSAGVMS-----VEFVSHITRSPFAISTKPTWLE-1199
DB 913 GAIGAVLAVMVSMNSDVYFKVGLITIGLSAKNAILIVEF-----AKELWEQ 961
QY 1200 -RAKEATISMGSAVFAVAMTNLP---GILVL-----GLAKAQLIQIFFRNLNLTLL 1249
DB 962 HSLRDAIAEARLRPRITMTSMAPLIGVPLALASGAGAASQRAIGTGVIGMLSATFL 1021
QY 1250 GLLHGLVFLPV 1260
DB 1022 GVL-----FVPI 1028

RESULT 85
US-09-902-540-10725
; Sequence 10725, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 10725
; LENGTH: 863
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-10725

Query Match 1.7%; Score 120; DB 2; Length 863;
Best Local Similarity 19.9%; Pred. No. 0.14;
Matches 154; Conservative 120; Mismatches 328; Indels 172; Gaps 34;

QY 549 PVFPFLAIGYKGKDYSEBALIMTFSLNYPAGDPRLAQAKLWEE-AFLEENRAFQRRM 607
DB 172 PMWDNTEIG--KTKDYLER---LKGQIDAYS--QPGKVKLVEDYKLMGDKTIAYGF 222
QY 608 AQMGQVTFTAERSLEDEINRTTAEDLPATSYVIVFLYISALGYSWSRWVDSKAT 667
DB 223 TGSYKTTVDSDFAIESLEPVTL-----IALGSIFLITITIFFRKLAP 264
QY 668 LGLGGVAVVLGAVMAAGFFSYLGRSSILVQLVFPVLVSLVGADNIFVLEYORLPER 727
DB 265 TFLVLVIGTVGVIY-TLGT-FTYATVGLNMTISILGILLGIDYGHFVPR-TRLELG 321
QY 728 PQGEVPHIGRALGRVAPSMLLCSLSEACFFLGLATTPMPAVRTFALTSGLAVIDFLIQ 787
DB 322 ACKPYDVAIRDVAGNAGPAVAAVTAGSPFVLVSEBFGFSQFGLAGMTLILGLTL 381
QY 788 MSAPVALLSLDSKREASRLDVCCCKVQDEL-----PPPGQEGELLGPFQKAYAPFLIH 842
DB 382 FCWSAAILAL-----AGRINP--ELPQKLGVMKPPD--TNSATTGKELRIPKPMVL 430
QY 843 WITRGVLLLFALFGVLSYSMCHISV---GLDQELALPKDSYLLDYFLFNRYFEVGAP 899
DB 431 GISTAIVALI-----CAAAPWAGSEE-----PPKVELGFFERLK-----466
QY 900 VYFVTLTGVNFSESA-----GMAACSSAGCN---NFSFTQKIQYATEPPE-----Q 943
```

```
DB 467 -YGVS---FNYNTRALIPDGMSSVLLQDEINERFNISDPMIAYTTKDLDEAGVYRELQ 522
QY 944 SYLATIPASSWDDFDIMLTPSSCC-----RLYISGPNKDKFCPSTVNSLN 988
DB 523 NAHKPSIDQVVSIFTFVPAPETAANAKVLEBKAEQMOLEDEGFSVAALPEMQANAD 582
QY 989 CLKNCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGLAAY-STSVNLTSDQVVL 1045
DB 583 FFKVKLDAKPFDPVHGVA-----NYTAQFEN-LPSAKPENHGVLTYIASVDL-MDQ-- 633
QY 1046 ASRFMAYHKPLKNSQDYTEALRAA-----RELAANITADLRKVPCTDPAFEVFP 1094
DB 634 -----KOLKFS-DETKVIAAYTPGFKDQDAMDPKAPTVEKBPRAAGAT-----QLYA 680
QY 1095 YTTITNVFEOYLTIPEGLFMLSICLVPTFAVSCLLGLDLRSLNLLSIVMLVDT--1152
DB 681 RLARIVLWDGKVTVLTAIWILAMHP-----LDFRNVKALASVIPLGVGVM 728
QY 1153 -VGFMALWDISYNAVSLINLVSAGVMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSA 1211
DB 729 MLGIMALTGLRLNFMNIIILPILGFGVSHGLYLLHRFLEGTSPVALRSVGAAVA-SST 787
QY 1212 VPAGVAMTNLPGLIVLGLAKAQLIQIFFRNLNLTLLGLHGLVPLPVLISV 1265
DB 788 LTAVVAFAALLAAAHNGL-----RSMGLVACIGLITTLVVSFTVLAHV 830
```

```
RESULT 86
US-09-902-540-14015
; Sequence 14015, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14015
; LENGTH: 735
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-14015

Query Match 1.7%; Score 118.5; DB 2; Length 735;
Best Local Similarity 27.5%; Pred. No. 0.15;
Matches 53; Conservative 32; Mismatches 59; Indels 49; Gaps 10;

QY 1136 RSGLLN-LLSIVMLVDT-----VGFMALWDISYNAVSLINLVS-----1174
DB 150 RAGTLSALIAVGLIILATERRVAALAVGIPVMGVVLTFAAAQLAIGHLNIYVTFELVAL 209
QY 1175 VQMSVEFVSHITRSPFAISTKPTWLER-----AKEATISMGSAVFAVGY---AMTNLPGLIVL 1227
DB 210 IGLGIEYGVHLKCMRY-----WEERRAHAPAREALVTAVRGTFSGAVTSVNTAAAFVVL 262
QY 1228 GLAKAQLIQIFFRNLNLTLLGLHGL-VFLPVLISY-VGPDVNPALALEOKRAEEAVA 1285
DB 263 LLAQQAQFNQP-----GLLAGLGLVLLAVLATVAMGSPSL-LALAEURLRPARVOVA 310
QY 1286 AMVAVSCNHPSR 1298
DB 311 AEATAPOVSQPER 323
```

```
RESULT 87
US-09-252-991A-20611
; Sequence 20611, Application US/09252991A
```


Db 506 --LIPFKAQ---KSVSVLPMPOBERFVTVGCAFPALAKALQFLHKLISQITID---VFF 557
Qy 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGLATPMPAVRTF----- 772
Db 558 IDWER-----PK-----GKVLKAV-----EGEGVRSATVPVSIWRTYFVANEM 596
Qy 773 -----ALTSGLAVILDPELLQMSAFVALLSLS-----KQEQASRLDVCCKVPQEL 818
Db 597 NBIQTVRKINSFQVLTVL--PFLEVGFKNLALMDSSLSRNPSPYIAPYSCI----- 649
Qy 819 PPPQGEGLLGFQKAYAPFLHWTIRGCVLLPLALP-----GVSLSVMCHI 867
Db 650 -----LRYAVSAALMLAIGIIQVVFAYFERFIEDKIQFVDLCSMSNI 694
Qy 868 SVGLDQELALPKDSYLLDYFLNRYFEVGAPYI-----FVTTLGYNFSSEAGMAICSS 922
Db 695 SV-----PILLSKCF--GYIIGHRSVHGADTNMEEMNNLAKREA--ENLCSQ 738
Qy 923 AGC--NNFSFTQKIQVATEPPQSYLAIPASSWVDDFIDWLPSSCCRLYISGPNKDKFC 980
Db 739 RGLVPNTDQQTPEIAISNQ--RQYDRIHET-----LIRKNGPARLLSS----- 782
Qy 981 PSTVNSLNLCKNCMSITMGSVRPSVEOPHKYLPWFLNDRPNIKCPKGGLAAYSTSVNLTS 1040
Db 783 ASTFQOSIKAYHMKFLGSP---IDHVHKMDYFIKDKLLE----- 822
Qy 1041 DQQVLASRFMAYHKPLKNSQDYTE 1064
Db 823 --RILGMEFM---EPMEKSIFYND 841

RESULT 89

US-09-902-540-11066
; Sequence 11066, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 11066
; LENGTH: 651
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-11066

Query Match 1.7%; Score 115; DB 2; Length 651;
Best Local Similarity 19.8%; Pred. No. 0.25;
Matches 129; Conservative 90; Mismatches 228; Indels 206; Gaps 32;
Qy 821 PGQEGLLIGFFQKAYAPFLHWTIRGVLL-----LFLAFGVSLYSNCHISVGLDQEL 875
Db 35 PAQME-----GYTKMIVNMML--FTNGIKNIETKNSQGLTMLKLSFYE-----GTDMGQ 82
Qy 876 ALPKDSYLLDYFLNLR-----YFEGAPVYFVTTLGYNFSSEAGMAICSSAGCNSFSFTQ 932
Db 83 AVAEINAL-----SNRSQVFLPPGAPPPFI--IRFDASSHPVGHVLPSES-----KTTN 130
Qy 933 KIQYATEPPEQSYL-AIPASSWVDDF-----IDWLTP-----SSCCRLY 970
Db 131 QLQDIANFTARPFLISIPGLTAPPGGSPRTIEINIEPEKLRVHNLTPQVVEIARON 190
Qy 971 ISGPNKDKFCPTVNSLNLCKNCMSITMGSVRPSVEOPHKYLPWFLNDRPNIKCPKGGLA 1030
Db 191 VTAPSGNVHGVGVD-----TYITPTNSTLR-AVEDF-----GNIPLLKGSVA 230

Qy 1031 -AYSTSVNLTSQGVLASRFMAYHKPLKNSQD--YTEALRAARELAANITADLRKVPQTD 1087
Db 231 NVYVRDVATVXGADIATGY-----ALVNGKRSVYLVNAKAGNASTVHVVVQQLK----- 279
Qy 1088 PAFEVPPYITIN-----VFYE-----QYLTILPEGLFMLSCLCLVPTTFAVSCLLGLDLRS 1137
Db 280 ---ESIPRIQSNLPDDVHSYBFDQSVYVVGALKGLIAEGVIGAVLTGLMVLFLRDVRS 336
Qy 1138 GLLNLLSLVIMILVDTVGFPMALWDISYNAVSLNLVSAVQMSVBFVSHITRSPASTKPTW 1197
Db 337 AIIVITTIPIAISGVLFKLFQGTINIMTLAGLALAGILVD-----ESTVTIENIHQH 391
Qy 1198 LERAKATISMGA-----VFA--GVAMTNLPGLIVLGLAKA---OLIQ 1236
Db 392 LQRGKTAVAVMDACLEIAFAFSKLLILLCLTSVFAPALTMGGIPGALFRPLALAIGCMVV 451
Qy 1237 IPFFRLNLLITLL-----GLL-----HGLVP 1257
Db 452 SFLLSQSLVPMANMMLKQHPGPGEGASPSGRRVRVIEGMLPHRKLILLVSAVGGLIV 511
Qy 1258 LPVI--LSYVGPDPNPAI-----ALEQKR-----AEEAVAA---VM 1288
Db 512 LAVVSLQRIGKDVLRVNSSQLRLRAAEGTRIEKTEQVHVQAMGLVEEVVCGAKVRIS 571
Qy 1289 VASCPNHPRSVSTANIYVNH-----SFECSIKGAGAISNPLPNRGQ 1331
Db 572 SAYVQGHPSFAISP-IYLYNAGPHEALLQVAFEGVGVDIDALKERIRQVRE 623

RESULT 90

US-09-902-540-12414
; Sequence 12414, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12414
; LENGTH: 810
; TYPE: PRF
; ORGANISM: Myxococcus xanthus
US-09-902-540-12414

Query Match 1.7%; Score 115; DB 2; Length 810;
Best Local Similarity 20.0%; Pred. No. 0.37;
Matches 169; Conservative 102; Mismatches 301; Indels 274; Gaps 37;
Qy 551 PFPLAIGGYKGYDSEAEA-----LIMITFSLNNYPAGDPR 585
Db 124 FATLGVSPVIGRAFTAEAVQGBRERVVLTDTAWRVHFPARDAQVLGQTMQLD---GEY 179
Qy 586 LAQAKL-----WEBAFL-----EEMRAFORRMAGMPQVTTFAERSLEDEINTTAED 632
Db 180 TVVGVLPAGVAYPAWAEMYVPVPTETQREQDQRGARFLSVVGRLLKPGVTDQAART---D 236
Qy 633 LPIFATSVIVIFLYISLALGS-YS---SWSRVWVDSKATLGLGVAVVGLGAVMAAGP 687
Db 237 LARVA-----LEMEBAVPSRYKRVKWSFVSTLDEKV---VGNVRGTLMILLGAVGV 286
Qy 688 SYLGRSSIVLIQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVHIGRAL-----GRV 743
Db 287 LLAACSSVANLL-----LARMARGREVSIRALGASRGVL 322
Qy 744 AFSMLLCSL-----SRAICFFLGALTMPA-----VNTFALTSLGLAVI 781

```
Db 323 VAQFLTESLVLSAGGVGLMLLALWGTDALVALVGEALPRASQVRLDARPLVFTMGVSL 382
Qy 782 LDFLQMSAFVALLSLDSKQASRLDVCCCVKQOELPPPGQGGELLGFFQKAYAPELL 841
Db 383 TGVLPGLGPA-----HGSREDLSVAMREGSRGTEGHGSG-----RMRAGLV 425
Qy 842 HWITGVVLLFLAFGVSLYSMCHISVGLDQ-----LALPKDSYLLDFLFLNRYFE 895
Db 426 GOVAVALLVGLGFTKSFALRAVDAGFTPEGVLTKGLVLP-----ARYPD 474
Qy 896 VGAPYVFTVTLGYNFSSBAGMAI-----CSSAGCNFSTFKIQIYATEFPFQSYLAIPA 950
Db 475 AAKHVAFORELLGRLOSLPGVEAVGVNLLPLGGRSDSP--DIEGRPOADVEWPVAF 532
Qy 951 SSWVDDFDWL-----PSSCCRLYISGPNKDKFCPSTVNSLCLK--- 991
Db 533 RVSVPGLRTRLVTPREGHLLGEGDTEGPDAPQVQVINKTFADLYWPOGDALGRLKLHW 592
Qy 992 -NCMSITMGSVRPSVEQPHKYLFWELND--RENIKCPKGLAAYSTSVNLTSDGOVLASR 1048
Db 593 DTAQWTTVVGIYDDVRE-----WGLDTFARP-----AAYTPAAKVPTFFALAVR 637
Qy 1049 FMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVFPYITINVFYEQYLT 1108
Db 638 AKSGSP-----EALRTA-----IEALRAVDGNLALYAVAPLT-----RL 673
Qy 1109 LPEGLFMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGFMAWMDISYNAVSL 1168
Db 674 VDES-----ISSRQNSALLMGLFAGTALL-----LAA 700
Qy 1169 INLVSAGMSVEFVSHITSPFAISTKPTWLERAKEATISMSGSAVPAGVAMTNLPGILVLG 1228
Db 701 LGISGVIGYS--VAQRTREMG I -RMALGASRVLTLVLGQ-----GLRLTALGVVLGLG 752
Qy 1229 LA--RAQLTQIF-----FFRLNLLITLLGLLHGLVFLPVLISYVGPVNVNPAAL 1275
Db 753 LSLGLARLLDAMLYGVAAYDGTWTFAGVALLGTVAVL--AAMP-----ARRATRVDPAL 807
Qy 1276 EOKRAE 1281
Db 808 ---RAE 810

RESULT 91
US-08-311-731A-363
; Sequence 363, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
```

```
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 363:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
; US-08-311-731A-363

Query Match 1.7%, Score 114; DB 2; Length 976;
Best Local Similarity 18.3%; Pred. No. 0.65;
Matches 151; Conservative 112; Mismatches 255; Indels 306; Gaps 37;

Qy 175 VRVPAATLAVGTMCVGVYGSALCNAQRWLN-----FQGDGTGNGLAPLDITFHLL 224
Db 273 VNLTSLAIAAGTDYGFITGRYQEARQANENKEAFTMYRG-----TFHVI- 320
Qy 225 PQQAVGSGIQPLNEGVARCNESQGDVATCSCQDCQCAASCPAIPARQALDSTFYLGOMPS 284
Db 321 ---LGSG-----TIS---GATECLSFARMPYFQT---LG-VPCA 350
Qy 285 LVLIILCSFANVTILLVGRVAPARPKSKMDPKKTSLSDKLSPTSTHLLGQFFQGW 344
Db 351 VGMLT-----AVAVALTGLPVLTVGSRFLGFEPRKLIK-----RGM 388
Qy 345 ---GTWVASWPLTILVSVIPVVALAAGLVFTTDPVELWSAPNS-QARSEKAFHDOH 400
Db 389 RRGTVVVRWPLPILITTC--AIAMVGLLALPGYRTNYKDRAYLPASIPANGQFAADRH 446
Qy 401 FGPPFRTNQVIL-----LQYFQNNRTLL-----LTA 505
Db 447 P-PQARMKEPILMTESDHDMENPADFLIDKLARGIFRVPGISRVAQITRPDGTAMDHTS 505
Qy 420 Y-----RYDSLILGPKNFS-GILDLDDLLLELLE-LQERLRLHQLW 457
Db 506 IPPFIQSMQVAGQVQTMKYQKDRMNDLLRQAEENMAETIASMRMRHQLMALLTHTHILND 565
Qy 458 SPEAORNIS-LQD-----ICVAPLNPNSTSLYDCC----- 486
Db 566 TVEMQRTSKLRDEANFDDFWRPISRYFYWERHCFNIPICWS-----FRSIFDALDGDV 620
Qy 487 ---INS-----LLQYFQNNRTLL-----LTA 505
Db 621 QIDERLNSIVGDIKNMDLLPQMLEQFPFPMIESMESMRTIMLTHTSTMSGIFDQMNELSD 680
Qy 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKDGITALALASCHADYCAPVPPPLAIGSYKGDYS 565
Db 681 NANTMGKAPDTAKNDDSFYL--PPEVFKN-----TDFKRAMKSPFUSSDGH----- 723
Qy 566 EAEALIMTFSLNYPAGDP-----RLAQKLWEEAFLEEMRAFORRMAGMQV 613
Db 724 AARFIL-----HRGDPASVAGTASINAIHTAAEEALKGTPLEDTKIYLAGTAAVFK- 775
Qy 614 TFTAERSLEDEINRTTAEDLPITAFSYIVIFLYISLALGSYSWSRWMDSKATLGLGV 673
Db 776 -----DIDEGANWDLVIAGISSLCLIFIIMLI-----TRAFV--AAAVIVGT 817
Qy 674 AVVLGAV--MAAMGFPSYLGIRSSVLQVVPFLVSLVGADNIFIFVLEYQRLPRRPGEP 731
Db 818 ALSLGASFGSLVLLWQHILGIELHVLVLAWSIVLVLAAGSDYNLLVLSRFK----- 868
Qy 732 REVHIG-----RALGRVAPSMLLCSLSSEACFFLGCALTMPFAVRTFALTSLGLVILDF 785
Db 869 QEIQAGLKTGIIRSMGTCVVTNAGLVFAFTMASVMVSDLRVIGQVGTIGLGLLFDTL 928
Qy 786 L-----QMSAFVALLSLDSR-----QASRLDVCCCVKQOELPP 820
```

```
Db 929 IYRSFMPMSIAALLG-----RWFHWPOQGRTRPLLTVAAPVGLPP 968

RESULT 92
US-09-252-991A-24417
; Sequence 24417, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24417
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24417

Query Match 1.7%; Score 114; DB 2; Length 1019;
Best Local Similarity 19.8%; Pred. No. 0.7;
Matches 162; Conservative 129; Mismatches 309; Indels 218; Gaps 43;

QY 580 PAGDP-----RLAQAKLWEAPLEEMRAFQRMAGHQPQVFTFAERSLEDEINRTTAEADIPI 635
Db 289 PSANPLDVIKYRAKMP-----LEEQLPNLKVSIAYDATRFIQASI-DEVVKTLGE----- 340
QY 636 FATSIVIVPLTSLAGSYSSSRVSDSKATLGLGVAVLGVAAAGPSPSYLGIRSS 695
Db 341 ---AVLIVVVVFLGAP-----RSVLIPVTPISMGVLP--FWQMGY-----S 383
QY 696 LVILQVVPFLVLSVG--ADNIFIFVLEYQRLPRPCEPREVHIGRALGRVAPSMLLCSLS 753
Db 384 INLLTLA-MVLAIGLVDDAIVVENIHRHTEEGKPPPEAALGAREIAVPVMSWITL 442
QY 754 EAICFPLGNLTMPAV--RTFALTSGLAVIDLFDLOMSAFVALLSDSKQF-----ASR 806
Db 443 AAVYAPIGFLTGLTALPKFEPFATAGAVIISGIVALTSLSPMNCSSRLRHEENPSGLAHR 502
QY 807 LDVCCVXPQELPPPGQGGILLGFPKAYAPFLHMI--TRGVVL---LLFLALFGVSL 861
Db 503 LDLIF-----EGLKQR-YQRA-----LHGTLDTRPVVLVFAVLVLAIPVLL 543
QY 862 YSMCHISVGLDQELALPKDSYL-----LDYFLFLNRY---PE---VGAPVVF- 902
Db 544 -----MFTKELAPEEDQGIVFLMTNSPQTANLDY---LNRYTAETEGIFRSFPEYYS 593
QY 903 -VTTGYNFPSSBAGNAICSSAGCNNSFTQKIQA-----TEFPQSYLAIPASSWDD 956
Db 594 AFQINGYN-GVQAGIGCMLLKPWDEREKSQEMELLHAQKLEIPGVQIFAFNLPS----- 648
QY 957 FIDMLTPSSCCRLYISGPNKPCFSTVNSLNCNCHSITWGSVRPSVQPHKYLPLWFL 1016
Db 649 -----LPGTGEGLPFPQVLTANDYESLIQVAQ-RVKQRASESQKFA--FL 691
QY 1017 N-----DRPNI-----KCPKGGGLAAVSTSVNLTs---DGQVLASRFW----- 1050
Db 692 DLDLADFPELVVDIDREKAAQMGVSMQDLGVALASLLGEGEI--NRTIDGRSVKYVIAQ 749
QY 1051 -----AVHKPLKNSQ-----DYTEALRAARELANIATADLRKVPGTDPAPFV 1092
Db 750 VERPYRDNBGLWYGVKSRNQALVALSTLIETHERARPRQLNQFQOLNSA-----IISG 804
QY 1093 RPY-----TITVVFYEQYLITLPEGL-----FMLSCLVPTFAVSCLLGL 1133
Db 805 PPIVSMGEAIEYVQOIAREE-----APRGFANDYAGASRQYVQEGSALLVTFGLALAIIFL 860

QY 1134 DL-----RSGLLNLLSIVMLVDTVGFMALWDISYNAVSLINLVSAGV-----MSV 1179
Db 861 VLAQOFESFRDPLVIMVTVPLSICGALIPFLFGVSSLSNIYTVQGLVTLGLSKHGILIV 920
QY 1180 BFVSHITRSFALSTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFP 1239
Db 921 EFANQLRHEQGIGRREA-IEQAAAIRLRPVLMTTAAWLVGLVPLILATGAG-----AVSR 974
QY 1240 FRANLLITL---LGLLHGLVFLFVLVSYG-EDVNPAL 1273
Db 975 PDGIVATGMSVGLFTLFTLVLPCTYTLVARDPGV 1012

RESULT 93
US-09-603-208A-226
; Sequence 226, Application US/09603208A
; Patent No. 6822084
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: BGI-124CP
; CURRENT APPLICATION NUMBER: US/09/603,208A
; CURRENT FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932914.1
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: DE 19940764.9
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19941382.7
; PRIOR FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 226
; LENGTH: 697
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-603-208A-226

Query Match 1.6%; Score 112.5; DB 2; Length 697;
Best Local Similarity 18.8%; Pred. No. 0.5;
Matches 183; Conservative 111; Mismatches 311; Indels 369; Gaps 44;

QY 341 FQGMGTWASWPLTTLVLVSIPVVALAAGL-----VFTBLTDPV--ELWSPNSQARSK 394
Db 2 FSKGHGFAFR-----RRIVPLVVAIALALFVIFGTGKGLGRMSQEGWDDPGSSSTAA 55
QY 395 AFHDQHFPPFTNQ-----VILTAPNRSRYDSLLGLPKPFGSGLDLDLLELELQER 450
Db 56 RIELETFG---RDNDGDVLLFTAPEGTSFDARVYF---SSISGYLD-----GLIE---- 100
```


216	L	S	G	-----	LSVAGVVMKLA	A	A	A	F	K	L	S	A	P	F	H	S	V	D	S	V	T	E	G	N	A	P	I	A	A	F	L	A	S	V	S	K	264																	
702	V	P	E	L	V	L	S	-----	DNIFIV	-----	LE	Y	O	R	L	P	R	R	P	G	E	P	R	E	V	H	I	G	737																										
265	V	A	M	M	A	L	R	F	F	G	S	M	T	A	M	P	E	T	I	L	I	V	I	V	L	S	I	L	M	G	N	L	L	A	I	R	324																		
738	R	A	L	G	R	V	-----	A	S	M	L	L	S	L	S	E	A	I	C	F	F	L	G	A	L	T	P	M	A	V	R	T	E	A	L	T	S	793																	
325	Y	A	L	T	A	L	S	T	G	A	S	L	P	L	V	S	M	T	A	V	-----	Y	A	L	T	S	-----	I	G	A	P	V	358																						
794	L	L	S	L	S	K	R	O	E	A	G	R	L	D	V	C	C	V	P	Q	E	L	P	P	P	G	O	G	E	L	L	G	F	F	Q	K	A	P	F	L	L	H	-----	I	T	R	G	V	L	850					
359	I	T	L	M	S	P	Y	R	E	R	G	T	S	-----	S	G	E	A	D	L	R	F	Y	Q	-----	L	F	W	R	R	P	V	L	T	A	V	L	399																	
851	L	L	E	F	A	L	F	G	V	S	L	S	M	C	H	I	S	V	G	L	D	O	E	L	A	L	P	K	D	S	V	L	D	D	Y	F	L	N	R	Y	E	V	E	G	A	P	V	F	V	T	I	G	N	F	910
400	V	M	L	S	L	A	G	I	P	L	-----	T	A	G	-----	F	V	T	K	F	-----	T	I	F	F	A	N	O	G	M	H	F	430																						
911	S	S	E	A	G	N	A	I	C	S	S	A	G	C	N	F	-----	S	F	T	O	K	I	Q	I	A	T	E	F	F	P	Q	S	Y	L	A	P	A	950																
431	L	A	-	A	G	M	I	L	G	S	A	I	G	L	V	Y	L	I	R	L	V	N	I	O	K	P	K	N	E	F	D	A	H	O	G	V	K	A	474																

```

RESULT 98
US-09-711-164-443
; Sequence 443, Application US/09711164
; Patent No. 6589738
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Alllyn
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyaskind, Judith
; TITLE OF INVENTION: GENES ESSENTIAL FOR MICROBIAL PROLIFERATION AND ANTISENSE THERAPY
; FILE REFERENCE: ELITRA.008A
; CURRENT APPLICATION NUMBER: US/09/711.164
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/164415
; PRIOR FILING DATE: 1999-11-9
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 443
; LENGTH: 1025
; TYPE: PR1
; ORGANISM: Escherichia coli
US-09-711-164-443

```

	Query Match	1.6%;	Score 111;	DB 2;	Length 1025;
	Best Local Similarity	22.4%;	Pred. No. 1.4;		
	Matches	71;	Conservative	50;	Mismatches 148; Indels 48; Gaps 12;
Qy	996	ITMGSVRPSVEOPFKYLP-----WFLNDRPNIKCP-----KGLAAYST	1034		
	:::	:::	:	:	:::
Db	199	VSLDDVRTAVSNANVRPQGGALEDGTHRWQIQTNDELKTAAEYOPLIIHYNNGSAGVRLGD	258		
Qy	1035	SVNLTSDSGVQLASRFMAHYHKPKNSQDYTEALRAAREL--AANTTADLRKVPGTDPAP--EV	1092		
	:::	:::	:	:	:::
Db	259	VATVTDSDVDVRNAGMTNAKB-----AILLMIRKLPEANIQTVDSTRAKLPELOET	310		
Qy	1093	FPHYITNNFYBOYLTIIPLEGFMLSCLIVPTFAVSCILLGLDURSG---LLNLLSIWMIL	1149		
	:::	:::	:	:	:::
Db	311	IIPAALDIQIAQDRSPTRASLEEEVEQTLIISVALVILWFELRSGRATIIPAVSVFVS	370		
Qy	1150	VDTVGFMALWDIISYNAVSLINLVSAVGMSVE----FVSHITRSPAISTKTWTWLERAKEAT	1205		
	:::	:::	:	:	:::
Db	371	IGTFAMMYLCGFSLNNLSLMALTATTGVDDDAIWLIENTARHIIEAGMKP--LQAALOGT	428		
Qy	1206	ISMGSAPVA---GVAMTNLPGLVLGLAKAQOLIQIFFFRNLMLITLLGLLHLGLFWLPVIL	1262		
	:::	:::	:	:	:::
Db	429	REVGETVLSMSSLVAVFLP-LLLMGGLPGRIILREPAVTLSVAIG-LSLLAVSLTLTPMC	486		
Qy	1263	SYVGPDVNPALALEOKR	1279		
	:::	:::	:	:	:::
Db	487	GWMLKASKPR---EOKR	500		

```

RESULT 99
US-09-492-709A-283
; Sequence 283, Application US/09492709A
; Patent No. 6720139
; GENERAL INFORMATION:
; APPLICANT: Zyskind, Judith
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Trawick, John
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Froelich, Jamie M.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: GENES IDENTIFIED AS REQUIRED FOR PROLIFERATION IN
; TITLE OF INVENTION: ESCHERICHIA COLI
; FILE REFERENCE: ELITRA.001A
; CURRENT APPLICATION NUMBER: US/09/492,709A
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 485
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 283
; LENGTH: 1025
; TYPE: PRT
; ORGANISM: E. Coli
US-09-492-709A-283

Query Match          1.6%; Score 111; DB 2; Length 1025;
Best Local Similarity 22.4%; Pred. No. 1.4;
Matches 71; Conservative 50; Mismatches 148; Indels 48; Gaps 12;

Qy 996 ITGSGVRPSVEQFHKYL P-----WFLNDRPNIKCP-----KGGLAAYST 1034
Db 199 VSLDDVRTASVNAVVRKPOGALEDGTHRWQIQTNDELKTAAYQPLIIHYNNGGAVRLGD 258
Qy 1035 SVNLTSQGVLAGRFMAYHKPLKNSQDYETALRAAREL--AANTADLRKVPGTDPAF--EV 1092
Db 259 VATVDSQVDVRAGMTNAP-----AILMIRKLPEANLIQTVDISRAKLPELQET 310
Qy 1093 FPYITITNVFEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLDRSG---LLNLLSVIMIL 1149
Db 311 IPAAIDLIQAQDRSPTRIASLEEVEQTIIISVALVILVWFLFIRSGRATIIPAVSVPVSL 370
Qy 1150 VDTVGFPMALWDISNVAVSLINLVSAGCMVSE----FVSHITRSPFAISTKPTMLERAKEAT 1205
Db 371 IGTFAAYLTCGFSLNLLSMALTATGTVVDDDAIVLNTARHLEAGMKP--LQALQGT 428
Qy 1206 ISMGSAVFA---GVAMTNLPGIIVLGLAKAQLIQIPFRLNLLITLGLLHGLVFLPVL 1262
Db 429 REVGFTVLMSLSLVAVFLP--LLLMGGLPGRLAREFAVTLVSAIG--ISLLAVSLTLTPMC 486
Qy 1263 SYVGDFVNPALAEQKR 1279
Db 487 GWMLKASKPR---EOKR 500

```

```

RESULT 100
US-09-489-039A-7216
; Sequence 7216, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489.039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7216
; LENGTH: 1048
; TYPE: PRT

```


GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:56:52 ; Search time 189 Seconds
(without alignments)
2944.702 Million cell updates/sec

Title: US-10-736-769-4
Perfect score: 6909
Sequence: 1 MAAEGLRGWLLWALLRLAQ.....GSIKGAIGAINFLPNNGRQF 1332

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : Published Applications AA.Main:*
1: /cgn2_6/prodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	4	US-10-621-758A-4
2	6909	100.0	1332	4	US-10-663-208A-4
3	6909	100.0	1332	4	US-10-646-301A-4
4	6909	100.0	1332	4	US-10-736-769-4
5	6896	99.8	1332	4	US-10-239-316-9
6	6872.5	99.5	1359	4	US-10-621-758A-4
7	6872.5	99.5	1359	4	US-10-663-208A-4
8	6872.5	99.5	1359	4	US-10-646-301A-4
9	6872.5	99.5	1359	4	US-10-736-769-4
10	6536	94.6	1344	5	US-10-450-763-53052
11	5421.5	78.5	1331	4	US-10-621-758A-2
12	5421.5	78.5	1331	4	US-10-663-208A-2
13	5421.5	78.5	1331	4	US-10-646-301A-2
14	5421.5	78.5	1331	4	US-10-736-769-2
15	5407	78.3	1333	4	US-10-621-758A-12
16	5407	78.3	1333	4	US-10-663-208A-12
17	5407	78.3	1333	4	US-10-646-301A-12
18	5407	78.3	1333	4	US-10-736-769-12
19	4466	64.6	982	5	US-10-450-763-53050
20	2402.5	34.8	1278	4	US-10-208-731-2
21	2402.5	34.8	1278	4	US-10-741-601-530
22	2402.5	34.8	1278	5	US-10-741-600-1542
23	2402.5	34.8	1278	5	US-10-756-149-4924
24	2365	34.5	1319	4	US-10-208-731-4
25	2262	32.7	1287	6	US-11-097-143-12003
26	1864.5	27.0	1223	6	US-11-097-143-2679
27	1329	19.2	1170	4	US-10-208-731-6

28	1078	15.6	1296	4	US-10-208-731-9	Sequence 9, Appli
29	942	13.6	194	4	US-10-239-316-8	Sequence 8, Appli
30	900.5	13.0	492	4	US-10-424-599-211862	Sequence 211862,
31	859.5	12.4	541	4	US-10-424-599-189288	Sequence 189288,
32	651.5	9.4	1447	2	US-08-954-701A-19	Sequence 19, Appli
33	651.5	9.4	1447	3	US-09-898-533-5	Sequence 5, Appli
34	651.5	9.4	1447	3	US-09-754-032-19	Sequence 19, Appli
35	651.5	9.4	1447	4	US-10-421-446-19	Sequence 19, Appli
36	651.5	9.4	1447	4	US-10-791-844-6	Sequence 6, Appli
37	646.5	9.4	1434	2	US-08-954-701A-10	Sequence 10, Appli
38	646.5	9.4	1434	3	US-09-754-032-10	Sequence 10, Appli
39	646.5	9.4	1434	4	US-10-421-446-10	Sequence 10, Appli
40	628	9.1	1296	4	US-10-302-279-60	Sequence 60, Appli
41	598.5	8.7	933	4	US-10-415-934-3	Sequence 3, Appli
42	597	8.6	821	5	US-10-890-776A-4805	Sequence 4805, Ap
43	588	8.5	954	4	US-10-060-756A-3	Sequence 3, Appli
44	588	8.5	954	5	US-10-890-776A-3	Sequence 3, Appli
45	584	8.5	1203	3	US-09-990-046-2	Sequence 2, Appli
46	583.5	8.4	1182	3	US-09-990-046-7	Sequence 7, Appli
47	582.5	8.4	343	4	US-10-425-115-328081	Sequence 328081,
48	580	8.4	1203	3	US-09-909-280A-2	Sequence 2, Appli
49	579	8.4	1146	4	US-10-791-844-1	Sequence 1, Appli
50	563	8.1	836	5	US-10-890-776A-4807	Sequence 4807, Ap
51	559.5	8.1	793	5	US-10-890-776A-4806	Sequence 4806, Ap
52	539.5	7.8	846	5	US-10-890-776A-4808	Sequence 4808, Ap
53	531	7.7	1061	6	US-10-415-934-9	Sequence 9, Appli
54	531	7.7	1061	6	US-11-097-143-23451	Sequence 23451, A
55	526.5	7.6	1311	2	US-08-954-701A-4	Sequence 4, Appli
56	526.5	7.6	1311	3	US-09-754-032-4	Sequence 4, Appli
57	526.5	7.6	1311	4	US-10-421-446-4	Sequence 4, Appli
58	514	7.4	1405	4	US-10-369-493-5464	Sequence 5464, Ap
59	474	6.9	1286	3	US-09-898-533-3	Sequence 3, Appli
60	474	6.9	1286	6	US-11-097-143-4068	Sequence 4068, Ap
61	464.5	6.7	1299	5	US-08-954-771-48	Sequence 48, Appli
62	464.5	6.7	1299	5	US-10-647-654-48	Sequence 48, Appli
63	464.5	6.7	1299	5	US-10-835-517-48	Sequence 48, Appli
64	442.5	6.4	906	5	US-10-890-776A-4809	Sequence 4809, Ap
65	438.5	6.3	1285	2	US-08-954-701A-6	Sequence 6, Appli
66	438.5	6.3	1285	3	US-09-754-032-6	Sequence 6, Appli
67	438.5	6.3	1285	4	US-10-421-446-6	Sequence 6, Appli
68	379	5.5	714	4	US-10-369-493-5410	Sequence 5410, Ap
69	351.5	5.1	767	4	US-10-060-756A-6	Sequence 6, Appli
70	351.5	5.1	767	5	US-10-890-776A-6	Sequence 6, Appli
71	339	4.9	783	4	US-10-094-749-1772	Sequence 1772, Ap
72	310.5	4.5	648	4	US-10-060-756A-10	Sequence 10, Appli
73	310.5	4.5	648	5	US-10-890-776A-10	Sequence 10, Appli
74	306.5	4.4	648	4	US-10-060-756A-4799	Sequence 4799, Ap
75	306.5	4.4	648	5	US-10-890-776A-4799	Sequence 4799, Ap
76	291	4.2	61	3	US-09-864-761-34826	Sequence 34826, A
77	283.5	4.1	136	4	US-10-425-115-277863	Sequence 277863,
78	272.5	3.9	132	4	US-10-425-115-277864	Sequence 277864,
79	266.5	3.9	792	4	US-10-415-934-8	Sequence 8, Appli
80	264.5	3.8	229	4	US-10-424-599-181755	Sequence 181755,
81	259	3.7	267	4	US-10-425-115-201210	Sequence 201210,
82	252	3.6	1505	4	US-10-408-765A-1316	Sequence 1316, Ap
83	245.5	3.6	5491	4	US-10-432-613-2	Sequence 2, Appli
84	230	3.3	785	4	US-10-415-934-2	Sequence 2, Appli
85	225	3.3	542	4	US-10-017-161-2024	Sequence 2024, Ap
86	225	3.3	542	4	US-10-292-798-1670	Sequence 1670, Ap
87	225	3.3	563	4	US-10-415-934-14	Sequence 14, Appli
88	214.5	3.1	1218	4	US-10-432-613-3	Sequence 3, Appli
89	214.5	3.1	1218	6	US-11-097-143-3123	Sequence 3123, Ap
90	212.5	3.1	1218	4	US-10-168-428-2	Sequence 2, Appli
91	210.5	3.0	1203	4	US-10-094-743-2651	Sequence 2651, Ap
92	202	2.9	1124	4	US-10-311-623-3	Sequence 3, Appli
93	197.5	2.9	1561	4	US-10-085-198-168	Sequence 168, App
94	195.5	2.8	1087	4	US-10-239-316-17	Sequence 17, Appli
95	195.5	2.8	1392	4	US-10-239-316-40	Sequence 40, Appli
96	190	2.8	100	4	US-10-437-963-102679	Sequence 102679,
97	187.5	2.7	931	4	US-10-168-428-3	Sequence 3, Appli
98	177	2.6	121	4	US-10-425-115-222635	Sequence 222635,
99	174	2.5	1176	4	US-10-041-018-267	Sequence 267, App
100	170	2.5	101	4	US-10-060-756A-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4
Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGRGWLLWALLRLAQSEPYTHIQPGYCAFYDECGKXPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGRGWLLWALLRLAQSEPYTHIQPGYCAFYDECGKXPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLLOKICPRLVTGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
DB 61 TPARKITGDHLLLOKICPRLVTGNTQACCSAKQLVSEASLSITKALLTRCPACSNF 120
QY 121 VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
DB 121 VNLHCHTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAFLDITFHLEPGQAVGSGIQPLNEGV 240
DB 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAFLDITFHLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESQGDVATCSQDCDCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESQGDVATCSQDCDCAAPARPOALDSTFYLGOMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSMVDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPFRTNQVILITAPNRSY 420
DB 361 IPVVALAAGLVTELTDPVELWSAPNSQARSEKAFHDQHPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNTSGIILDLLELLELQERLRLHQLWSPEAQRNLSLQDICVAPLNPDNT 480
DB 421 RYDLSLLGPKNTSGIILDLLELLELQERLRLHQLWSPEAQRNLSLQDICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFKDGFTAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYLCANAPLTFKDGFTAL 540
QY 541 SCHADYGAVFPFLAIGGKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
DB 541 SCHADYGAVFPFLAIGGKGYDSEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
QY 601 RAPQRMAGMFQVTTAERSLEDEINRTTAEDLPIPATSYIVIFLYISLALGSYSSWSRV 660

DB 601 RAPQRMAGMFQVTTAERSLEDEINRTTAEDLPIPATSYIVIFLYISLALGSYSSWSRV 660
QY 661 MVDKATILGLGGVAVVILGAVNAANGPFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MVDKATILGLGGVAVVILGAVNAANGPFSYLGIRSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHIGRALGRVAPSMLLCSLSAICFFLIGALTMPAVRTPALTSGLAV 780
DB 721 YORLPRRPGEPREVIHIGRALGRVAPSMLLCSLSAICFFLIGALTMPAVRTPALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVXPQELPPPPQSGGLLGGFPQKAYAPPL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVXPQELPPPPQSGGLLGGFPQKAYAPPL 840
QY 841 LHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVITLGYNFSSEAGMNAICSSAGCNSFTQKIQYATEFFPEQSYLAIPASSWVDDFIDW 960
DB 901 YFVITLGYNFSSEAGMNAICSSAGCNSFTQKIQYATEFFPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLXIISGPNKDKFCFSTVNSLNCNCKMSITMGSVRPSVEQFHKYLPWFINDRP 1020
DB 961 LTPSSCCRLXIISGPNKDKFCFSTVNSLNCNCKMSITMGSVRPSVEQFHKYLPWFINDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDPAFVFPYTIINNVFYEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFVFPYTIINNVFYEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
QY 1141 NLLSIVMLVDTVGPMAWDISYNVSLINLVSAVMSVPEVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIVMLVDTVGPMAWDISYNVSLINLVSAVMSVPEVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISGSAVFAVAMTNLPGIILVLGAKAQLIQIIFFRNLNLTITLLGLHGLVFLPV 1260
DB 1201 AKEATISGSAVFAVAMTNLPGIILVLGAKAQLIQIIFFRNLNLTITLLGLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSRVSTADNIVVNHSPGSIKAGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSRVSTADNIVVNHSPGSIKAGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 2
US-10-663-208A-4
; Sequence 4, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1

Db 181 ATLAGVTGCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILITAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILITAPNRSSY 420
Qy 421 RYDLSLLGPKNPSGILDLDDLLELLELQERLRLHQLQWSPEAQRNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLLELLELQERLRLHQLQWSPEAQRNISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDKMDHFLYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDKMDHFLYCANAPLTFKQGTALAL 540
Qy 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEBAFLSEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEBAFLSEM 600
Qy 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSWSRV 660
Db 601 RAFORMAGMFOVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSWSRV 660
Qy 661 MYDSKATILGLGVAVVLGAVMAAMGFESYLGIRSSILVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAMGFESYLGIRSSILVILQVVPFLVLSVGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMILCSLSEAI CFFLGALTPMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCVKQOEPLPPQGGGLLIGFPQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCCVKQOEPLPPQGGGLLIGFPQKAYAPFL 840
Qy 841 LHWITRGVLLAFALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
Db 841 LHWITRGVLLAFALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSFTQKIYATEFPFQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFFSFTQKIYATEFPFQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMSITMGSVRPSVEQPHKYLFWFLNDRP 1020
Qy 1021 NIKCPKGGAAVYSTSNLSDGOVLASRFMAVHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGGAAVYSTSNLSDGOVLASRFMAVHKPLKNSQDYTEALRAARELANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYITINVFYEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLSGGL 1140
Db 1081 RKVPGTDPAFEVFPYITINVFYEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLSGGL 1140
Qy 1141 NLLSIVMLVDTVGFMAWMDISVNASVLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFMAWMDISVNASVLINLSVAGMSVEFVSHITRSPAISTKPTWLER 1200
Qy 1201 AKEATISMSGSAVAGVAMTNLPGILVLGLAKAQILQIFPFRLLNLTLLGLLHGLVFLPV 1260
Db 1201 AKEATISMSGSAVAGVAMTNLPGILVLGLAKAQILQIFPFRLLNLTLLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPVNPALALBQKRAEEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFEGSIKAGA 1320
Db 1261 ILSYVGPVNPALALBQKRAEEAAVAVMVASCPNHPFSRVSTADNIYVNHSPFEGSIKAGA 1320

Qy 1321 ISNPLPNNGRQF 1332
Db 1321 ISNPLPNNGRQF 1332

RESULT 4

US-10-736-769-4
; Sequence 4, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J801603-K3-US
; CURRENT APPLICATION NUMBER: US/10736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-4

Query Match 100.0%; Score 6909; DB 4; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Qy 61 TPARKITGDHILLQKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILLQKICPRLYTGNTQACCSAKOLVSEASLSITKALLTRCPACSDNF 120
Qy 121 VNLCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAGVTGCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAGVTGCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSCQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVTI 300
Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILITAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILITAPNRSSY 420
Qy 421 RYDLSLLGPKNPSGILDLDDLLELLELQERLRLHQLQWSPEAQRNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNPSGILDLDDLLELLELQERLRLHQLQWSPEAQRNISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDKMDHFLYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSOVDKMDHFLYCANAPLTFKQGTALAL 540

Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTTKDGTTALAL	540
Qy	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLSEM	600
Db	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLSEM	600
Qy	601	RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV	660
Db	601	RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV	660
Qy	661	MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSGADNIRIFVLE	720
Db	661	MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSGADNIRIFVLE	720
Qy	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA	780
Db	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA	780
Qy	781	ILDFLQMSAFVALLSDSKRQESRLDVCCCKVQBELPPPGQEGLLIGFTQKAYAPFL	840
Db	781	ILDFLQMSAFVALLSDSKRQESRLDVCCCKVQBELPPPGQEGLLIGFTQKAYAPFL	840
Qy	841	LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Qy	901	YFVTTLYGNFSSEAGNNAICSSAGCNCFSTQKIQVATEFPQSYLAIPASSWVDVDFIDW	960
Db	901	YFVTTLYGNFSSEAGNNAICSSAGCNCFSTQKIQVATEFPQSYLAIPASSWVDVDFIDW	960
Qy	961	LTPSSCCRLYISGPNKDFCPSTVNSLCLKNQMSITMGSVRPSVEQFHXYLPWFINDRP	1020
Db	961	LTPSSCCRLYISGPNKDFCPSTVNSLCLKNQMSITMGSVRPSVEQFHXYLPWFINDRP	1020
Qy	1021	NIKCPKGGAAVYSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL	1080
Db	1021	NIKCPKGGAAVYSTSVNLTSDQVLASRFMAHYHKLKNSQDYTEALRAARELAANITADL	1080
Qy	1081	RKVPDTPAFEPFPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140
Db	1081	RKVPDTPAFEPFPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL	1140
Qy	1141	NLLSIWMLVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER	1200
Db	1141	NLLSIWMLVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLER	1200
Qy	1201	AKEATISMSGSAPFAGVAMTNLPGILVLGAKAQLIQIIPFRNLALITLLGLHGLVFLPV	1260
Db	1201	AKEATISMSGSAPFAGVAMTNLPGILVLGAKAQLIQIIPFRNLALITLLGLHGLVFLPV	1260
Qy	1261	ILSYVGPVNPALALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPSGSKGAGA	1320
Db	1261	ILSYVGPVNPALALEQKRAEAAVAVMVASCPNHPSPRSTADNIYVNHSPSGSKGAGA	1320
Qy	1321	ISNFLPNNGRQF 1332	
Db	1321	ISNFLPNNGRQF 1332	

RESULT 5
US-10-239-316-9
; Sequence 9, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANITAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1e1 Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22

; PRIOR APPLICATION NUMBER: JP2000-088595 ; PRIOR FILING DATE: 2000-03-24 ; NUMBER OF SEQ ID NOS: 59 ; SEQ ID NO 9 ; LENGTH: 1332 ; TYPE: PRT ; ORGANISM: Human ; US-10-239-316-9			
Query Match 99.8%; Score 6896; DB 4; Length 1332; Best Local Similarity 99.8%; Pred No. 0; Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECKNPELSGLMTLSNVSCLSN	60
Db	1	MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECKNPELSGLMTLSNVSCLSN	60
Qy	61	TPARKITGDHLILLOKICPRLYTGNTOACCSAKQLVSLBASLSITKALLTRCPACSDNF	120
Db	61	TPARKITGDHLILLOKICPRLYTGNTOACCSAKQLVSLBASLSITKALLTRCPACSDNF	120
Qy	121	VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAEAFYQHSFARQSYDSCSRVRPAA	180
Db	121	VNLHCHNTCSPNOSLFINVTRVAQLGAGOLPAVVAEAFYQHSFARQSYDSCSRVRPAA	180
Qy	181	ATLAVGTMCVGYGSALCNAQRWLNFGDGTENGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240
Db	181	ATLAVGTMCVGYGSALCNAQRWLNFGDGTENGLAPLDITFHLLEPGQAVGSGIQPLNEGV	240
Qy	241	ARCNSQGGDVATCQDCCAASCPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Db	241	ARCNSQGGDVATCQDCCAASCPAIARQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300
Qy	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSV	360
Db	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSV	360
Qy	361	IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHDOHGFPPRTNQTIVITAPNRSY	420
Db	361	IPVVALAAGLVTELTTPDELWSAPNSQARSEKAFHDOHGFPPRTNQTIVITAPNRSY	420
Qy	421	RYDSLGLLPKNPSGILDLLELLELQERLHLQVMSPEAQRNLSQDICVAPLNDPT	480
Db	421	RYDSLGLLPKNPSGILDLLELLELQERLHLQVMSPEAQRNLSQDICVAPLNDPT	480
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTTKDGTTALAL	540
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTTKDGTTALAL	540
Qy	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLSEM	600
Db	541	SCMADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAQKLEAEFLSEM	600
Qy	601	RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV	660
Db	601	RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALAGSYSSWSRV	660
Qy	661	MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSGADNIRIFVLE	720
Db	661	MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSGADNIRIFVLE	720
Qy	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA	780
Db	721	YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSGLA	780
Qy	781	ILDFLQMSAFVALLSDSKRQESRLDVCCCKVQBELPPPGQEGLLIGFTQKAYAPFL	840
Db	781	ILDFLQMSAFVALLSDSKRQESRLDVCCCKVQBELPPPGQEGLLIGFTQKAYAPFL	840
Qy	841	LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900
Db	841	LHWITRGVLLFLALFGVLSYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900

```
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNCSITMGSVRPSVEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNCSITMGSVRPSVEQPHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 RKVPGTDAFEVFPYTTINVFYEQVLTILPEGLFMLSCLVPTPAVSCLLGLDLSGLL 1140
DB 1081 RKVPGTDAFEVFPYTTINVFYEQVLTILPEGLFMLSCLVPTPAVSCLLGLDLSGLL 1140
QY 1141 NLLSVIMILVDVTFGMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSVIMILVDVTFGMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMSGSAFVAGVAMTNLPGLILVLGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260
DB 1201 AKEATISMSGSAFVAGVAMTNLPGLILVLGLAKAQLIQIFFRLNLLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVNPALALEOKAEAEAAVAVMVASCPNHPSRVSTADNIIYVNHSPGSIKGAGA 1320
DB 1261 ILSYVGPVNPALALEOKAEAEAAVAVMVASCPNHPSRVSTADNIIYVNHSPGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 6
US-10-621-758A-44
; Sequence 44, Application US/10621758A
; Publication No. US20040093629A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-621-758A-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGRGWLLWALLRLAQSEPYTTTHORGYCAFYDECCKNPGLSGSLMTLSNVSCLSN 60
DB 1 MAEAGRGWLLWALLRLAQSEPYTTTHORGYCAFYDECCKNPGLSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVLEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSFNOSLFINVTRVAQLGAGQLPAVAVAEAFYQHSFABQS YDCSRVRVPA 180
DB 121 VNLHCHTCSFNOSLFINVTRVAQLGAGQLPAVAVAEAFYQHSFABQS YDCSRVRVPA 180
QY 181 ATLAAGTMCVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
```

```
DB 181 ATLAAGTMCVGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGIQPLNEGV 240
QY 241 ARCNESSODDVAATSCODCAASCPAIARPOALDSTFVLGOMPGSLVLIILICSVFVAVTI 300
DB 241 ARCNESSODDVAATSCODCAASCPAIARPOALDSTFVLGOMPGSLVLIILICSVFVAVTI 300
QY 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDOHGFPPFRFNQVTLTAPNRSSY 420
DB 361 IPVVALAAGLVFTLTTDPVELWSAPNSQARSEKAFHDOHGFPPFRFNQVTLTAPNRSSY 420
QY 421 RYDLSLLGPKNFSGILDLDLLELELOERLHIOVMSPEAORNISLQDI CYAPLNPNPT 480
DB 421 RYDLSLLGPKNFSGILDLDLLELELOERLHIOVMSPEAORNISLQDI CYAPLNPNPT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAALAL 540
QY 541 SCMADYGAPVPPFLAIGYKGYSEBALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCMADYGAPVPPFLAIGYKGYSEBALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFORMMAGMFOVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGYSWSRVR 660
DB 601 RAFORMMAGMFOVTFMAERSLEDEINRTTADLPIFATSYIVIFLYISLALGYSWSRVR 660
QY 661 MVDGKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVLQVVPFLVLSVGADNIFIVLE 720
DB 661 MVDGKATLGLGGVAVVLGAVMAAMGFFSYLGIRSSLVLQVVPFLVLSVGADNIFIVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLICALTPMPAVRFTALTSGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMLLCSLSEAI CFFLICALTPMPAVRFTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPOELPPGQGGELLLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEARLDVCCVKPOELPPGQGGELLLGFFQKAYAPFL 840
QY 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPV 900
DB 841 LHWITRGVVLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNCSITMGSVRPSVEQPHKYLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNCSITMGSVRPSVEQPHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDQVLAASRPMAYHKPLKNSQDYTEALRAARELANITADL 1080
QY 1081 KPLKNSQDYTEALRAARELANITADL RKVPGTDAFEVFPYTTINVFYEQVLTILPEGL 1140
DB 1081 KPLKNSQDYTEALRAARELANITADL RKVPGTDAFEVFPYTTINVFYEQVLTILPEGL 1140
QY 1141 FMLSCLVPTPAVSCLLGLDLSGLLNLISVIMILVDVTFGMALWDISYNAVSLINLVS 1200
DB 1141 FMLSCLVPTPAVSCLLGLDLSGLLNLISVIMILVDVTFGMALWDISYNAVSLINLVS 1200
QY 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMSGSAFVAGVAMTNLPGLILVLGLAKA 1260
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMSGSAFVAGVAMTNLPGLILVLGLAKA 1260
QY 1234 LQIIFPRNLNLLITLGLLHGLVFLPVILSYVGPVNPALALEOKAEAEAAVAVMVASCP 1293
DB 1234 LQIIFPRNLNLLITLGLLHGLVFLPVILSYVGPVNPALALEOKAEAEAAVAVMVASCP 1293
```

Db 1261 LIQIPPRNLITLLGLLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1320
QY 1294 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1359

RESULT 7

US-10-663-208A-44
; Sequence 44, Application US/10663208A
; Publication No. US20040132058A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-44

Query Match 99.5%; Score 6872.5; DB 4; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPGLSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAPYDECGKNPGLSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVSLASLSTIKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPLRYTGPNTOACCSAKQLVSLASLSTIKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAOLGAGOLPAVAYEAPYQHSFAEQSDSCSRVRVPA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAOLGAGOLPAVAYEAPYQHSFAEQSDSCSRVRVPA 180
QY 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNSQGDVATCSQDCAASCPAIAAPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGDVATCSQDCAASCPAIAAPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQPFQGWGTWASWPLTILVLSV 360
Db 301 LLVGFRVAPARDKSNVDPKGTSLSDKLSFSTHTLLGQPFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTDDVELWSAPNSQARSEKAFPHQFGPPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVTELTDDVELWSAPNSQARSEKAFPHQFGPPFRTNQVILTAPNRSY 420
QY 421 RYDLSLLGPNKPSGILDDLLELLELQERHLQVWSPQAQRNLSLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPNKPSGILDDLLELLELQERHLQVWSPQAQRNLSLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTPKDTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTPKDTALAL 540

QY 541 SCWADYGAPVPFPFLAIGGYKGDYSEAEALIMTFSILNNYPAGDPRLAQAKLWEEAEFL 600
Db 541 SCWADYGAPVPFPFLAIGGYKGDYSEAEALIMTFSILNNYPAGDPRLAQAKLWEEAEFL 600
QY 601 RAPQRMAGMFQVTTAERSLEDEINRTTAEPLDPATSYIIVIFLYISIALAGSYSSWSRV 660
Db 601 RAPQRMAGMFQVTTAERSLEDEINRTTAEPLDPATSYIIVIFLYISIALAGSYSSWSRV 660
QY 661 MVDSKATLGLGGVAVVLGAVMAAGFPFSSYLGIRSSILVILQVVPFLVSLGADNIRIFVLE 720
Db 661 MVDSKATLGLGGVAVVLGAVMAAGFPFSSYLGIRSSILVILQVVPFLVSLGADNIRIFVLE 720
QY 721 YQRLPRRFPCEPREVHIGRALGRVAPSMCLCSLSEAI CFFELGALTTPMPAVRTTALTSGLAV 780
Db 721 YQRLPRRFPCEPREVHIGRALGRVAPSMCLCSLSEAI CFFELGALTTPMPAVRTTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOEPLPPCGQEGLLIGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOEPLPPCGQEGLLIGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSFTQKIQVATEFPPEQSYLAIPASSWVDDFDW 960
Db 901 YFVTTLYGNFSSEAGMNAICSSAGCNPFSFTQKIQVATEFPPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCMLKNCMSITMGSVRPSVEQPHKYLPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCMLKNCMSITMGSVRPSVEQPHKYLPWFLNDRP 1020
QY 1021 NIKPKGGLAAVSTSVNLTSDQVLI-----ASRFMAVH 1053
Db 1021 NIKPKGGLAAVSTSVNLTSDQVLDVTVAILSPRLSEYSGTISAHNCNLYLLDSASRPMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVFPYTTITNVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELANITADLRKVPCTDPAFVFPYTTITNVFYEQYLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLNLLSITVILVDTVGPMAWDISYNAVSILNLS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRLSGLNLLSITVILVDTVGPMAWDISYNAVSILNLS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVPAVAMTNLPGIILVGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVPAVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIPPRNLITLLGLLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1293
Db 1261 LIQIPPRNLITLLGLLHGLVFLPVILSYGPDVNPALALEQKRAEAAVAVASCP 1320
QY 1294 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1359

RESULT 8

US-10-646-301A-44
; Sequence 44, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 10/621,758									
; PRIOR FILING DATE: 2003-07-17									
; NUMBER OF SEQ ID NOS: 50									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 44									
; LENGTH: 1359									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-646-301A-44									
Query Match 99.5%; Score 6872.5; DB 4; Length 1359;									
Best Local Similarity 97.9%; Pred. No. 0;									
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;									
Qy	1	MAEAGLRGMLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60						
Db	1	MAEAGLRGMLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60						
Qy	61	TPARKITGDHLILLOKICRLPYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120						
Db	61	TPARKITGDHLILLOKICRLPYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120						
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180						
Db	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180						
Qy	181	ATLAVGTMCVGYGSALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240						
Db	181	ATLAVGTMCVGYGSALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240						
Qy	241	ARCNSQGGDVATCQDCAACPAIRPAQALDSTFYLGOMPGSLVLIILCSVFVVTI	300						
Db	241	ARCNSQGGDVATCQDCAACPAIRPAQALDSTFYLGOMPGSLVLIILCSVFVVTI	300						
Qy	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV	360						
Db	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV	360						
Qy	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420						
Db	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420						
Qy	421	RYDSLGLGPNFSGIILDLLELLELQERLRLQWSPQARNISLQDICVAPLNPDNT	480						
Db	421	RYDSLGLGPNFSGIILDLLELLELQERLRLQWSPQARNISLQDICVAPLNPDNT	480						
Qy	481	SLYDCCINSLLQYFQNNRTLLILLTANQTLMGOTSQVDWKDHFYCANAPLTFKQGTALAL	540						
Db	481	SLYDCCINSLLQYFQNNRTLLILLTANQTLMGOTSQVDWKDHFYCANAPLTFKQGTALAL	540						
Qy	541	SCMADYGAPVFPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEAPLEEM	600						
Db	541	SCMADYGAPVFPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEAPLEEM	600						
Qy	601	RAFQRMAGNFQVTFMAERSLEINRTAEDLPIEATSVIVFLYISLALGYSWSRSV	660						
Db	601	RAFQRMAGNFQVTFMAERSLEINRTAEDLPIEATSVIVFLYISLALGYSWSRSV	660						
Qy	661	MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSGADNIFIVLE	720						
Db	661	MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSGADNIFIVLE	720						
Qy	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780						
Db	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780						
Qy	781	ILDFLLQMSAFVALLSLDSKQASRLDVCCKVQBELPPGQEGILLGFFQKAVAPFL	840						
Db	781	ILDFLLQMSAFVALLSLDSKQASRLDVCCKVQBELPPGQEGILLGFFQKAVAPFL	840						
Qy	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLADYFLFLNRYFEVGAPV	900						
Db	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLADYFLFLNRYFEVGAPV	900						
RESULT 9									
US-10-736-769-44									
; Sequence 44, Application US/10736769									
; Publication No. US20040161838A1									
; GENERAL INFORMATION: Scott W									
; APPLICANT: Altman, Scott W									
; APPLICANT: Wang, Luquan									
; APPLICANT: Graziano, Michael									
; APPLICANT: Murgolo, Nick									
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF									
; FILE REFERENCE: JB01603-K3-US									
; CURRENT APPLICATION NUMBER: US/10/736,769									
; CURRENT FILING DATE: 2003-12-16									
; PRIOR APPLICATION NUMBER: 60/397,442									
; PRIOR FILING DATE: 2002-07-19									
; PRIOR APPLICATION NUMBER: 10/621,758									
; PRIOR FILING DATE: 2003-07-17									
; PRIOR APPLICATION NUMBER: 10/646,301									
; PRIOR FILING DATE: 2003-08-22									
; PRIOR APPLICATION NUMBER: 10/663,208									
; PRIOR FILING DATE: 2003-09-16									
; NUMBER OF SEQ ID NOS: 51									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 44									
; LENGTH: 1359									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-736-769-44									
Query Match 99.5%; Score 6872.5; DB 4; Length 1359;									
Best Local Similarity 97.9%; Pred. No. 0;									
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;									
Qy	1	MAEAGLRGMLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60						
Db	1	MAEAGLRGMLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKNPELSGLMTLSNVCSLSN	60						
Qy	61	TPARKITGDHLILLOKICRLPYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120						
Db	61	TPARKITGDHLILLOKICRLPYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120						
Qy	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180						
Db	121	VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA	180						
Qy	181	ATLAVGTMCVGYGSALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240						
Db	181	ATLAVGTMCVGYGSALCAQRWLNFGDGTGNGLAPLDTFHLLPEQAVGSGIQPLNEGV	240						
Qy	241	ARCNSQGGDVATCQDCAACPAIRPAQALDSTFYLGOMPGSLVLIILCSVFVVTI	300						
Db	241	ARCNSQGGDVATCQDCAACPAIRPAQALDSTFYLGOMPGSLVLIILCSVFVVTI	300						
Qy	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV	360						
Db	301	LLVGRVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV	360						
Qy	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420						
Db	361	IPVVALAAGLVFTELTTDPVELWASPNQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420						
Qy	421	RYDSLGLGPNFSGIILDLLELLELQERLRLQWSPQARNISLQDICVAPLNPDNT	480						
Db	421	RYDSLGLGPNFSGIILDLLELLELQERLRLQWSPQARNISLQDICVAPLNPDNT	480						
Qy	481	SLYDCCINSLLQYFQNNRTLLILLTANQTLMGOTSQVDWKDHFYCANAPLTFKQGTALAL	540						
Db	481	SLYDCCINSLLQYFQNNRTLLILLTANQTLMGOTSQVDWKDHFYCANAPLTFKQGTALAL	540						
Qy	541	SCMADYGAPVFPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEAPLEEM	600						
Db	541	SCMADYGAPVFPFLAIGGYKGDYSBAALIMTFSLNYPAGDPRLAQAKLWEAPLEEM	600						
Qy	601	RAFQRMAGNFQVTFMAERSLEINRTAEDLPIEATSVIVFLYISLALGYSWSRSV	660						
Db	601	RAFQRMAGNFQVTFMAERSLEINRTAEDLPIEATSVIVFLYISLALGYSWSRSV	660						
Qy	661	MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSGADNIFIVLE	720						
Db	661	MVDSKATLGGVAVVLGAVMAAGPFSYLGIRSSILVILQVVPFLVLSGADNIFIVLE	720						
Qy	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780						
Db	721	YQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFLIGALTTPMPAVRTFALTSGLAV	780						
Qy	781	ILDFLLQMSAFVALLSLDSKQASRLDVCCKVQBELPPGQEGILLGFFQKAVAPFL	840						
Db	781	ILDFLLQMSAFVALLSLDSKQASRLDVCCKVQBELPPGQEGILLGFFQKAVAPFL	840						
Qy	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLADYFLFLNRYFEVGAPV	900						
Db	841	LHMTTRGVLLLFALFGVSLYSNCHISVGLDQELALPKDSYLLADYFLFLNRYFEVGAPV	900						

Db 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180
Qy 181 ATLAVCTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPQQAQVSGIQLNEGV 240
Db 181 ATLAVCTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPQQAQVSGIQLNEGV 240
Qy 241 ARCNEQSGDDVATCSQDCAASCAPAIARPOALDSTFYLGMQPSGLVLIILCSVFAVVTI 300
Db 241 ARCNEQSGDDVATCSQDCAASCAPAIARPOALDSTFYLGMQPSGLVLIILCSVFAVVTI 300
Qy 301 LLVGFVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWELTILVLSV 360
Db 301 LLVGFVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWELTILVLSV 360
Qy 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHDQHPGFPFRTNQVILITAPNRSY 420
Qy 421 RYDSSLGPKNFGSGLLDLLELLELQERLRHLQVWSPQARNISLODICYAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLLDLLELLELQERLRHLQVWSPQARNISLODICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFCQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFCQGTALAL 540
Qy 541 SCADYGAVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKWEAFLEEM 600
Db 541 SCADYGAVPFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKWEAFLEEM 600
Qy 601 RAPORMAGMFWQTFABRSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660
Db 601 RAPORMAGMFWQTFABRSLEDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660
Qy 661 MYDSKATLGLGVAVVLGAVMAAMGFSSYLGRSSILVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFSSYLGRSSILVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGVAPSMLLCSLSEAI CFFPLGALTMPAVRTPALTSGLA 780
Db 721 YQRLPRRPGEPREVIHGRALGVAPSMLLCSLSEAI CFFPLGALTMPAVRTPALTSGLA 780
Qy 781 ILDELLQNSAFVALLSLDSKQREASRLDYCCCKVQDELPPQCGEGLLGFPQKAYAPFL 840
Db 781 ILDELLQNSAFVALLSLDSKQREASRLDYCCCKVQDELPPQCGEGLLGFPQKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAVP 900
Db 841 LHWITRGVVLLFLALFGVSLYSMCHI SVGLQDELALPKDSYLLDYFLPLNRYFEVGAVP 900
Qy 901 YFVTTLGYNFSSEAGMNAICSSAGCNGNFSTQKI QYATEFPEQSYLAIPASSWVDFFIDW 960
Db 901 YFVTTLGYNFSSEAGMNAICSSAGCNGNFSTQKI QYATEFPEQSYLAIPASSWVDFFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSRVPSVEQFKHYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMSITMGSRVPSVEQFKHYLPWFINDRP 1020
Qy 1021 NIKCPKGGGLAAYSTSVNLTSDQVL-----ASRFMAVH 1053
Db 1021 NIKCPKGGGLAAYSTSVNLTSDQVL-----ASRFMAVH 1053
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1140
Qy 1114 FMLSICLVPTFAVSCLLGLDRSGLLNLSIVMILVDTVGPMAWLDISYNAVSLINLVS 1173
Db 1141 FMLSICLVPTFAVSCLLGLDRSGLLNLSIVMILVDTVGPMAWLDISYNAVSLINLVS 1200

RESULT 10

US-10-450-763-53052
; Sequence 53052, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)..(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX.
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052

Query Match 94.6%; Score 6536; DB 5; Length 1344;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

Qy 1 MAEAGRGWLLWALLRLAQSEPTTIHQPGYCAPYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGRGWLLWALLRLAQSEPTTIHQPGYCAPYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLQKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFABQSDSCSRVRPAA 180
Qy 181 ATLAVCTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPQQAQVSGIQLNEGV 240
Db 181 ATLAVCTMCGVYGSALCNAQRWLNFGQDGTGNGLAPLDITFHLLEPQQAQVSGIQLNEGV 240
Qy 241 ARCNEQSGDDVATCSQDCAASCAPAIARPOALDSTFYLGMQPSGLVLIILCSVFAV 298
Db 241 ARCNEQSGDDVATCSQDCAASCAPAIARPOALDSTFYLGMQPSGLVLIILCSVFAV 298
Qy 299 TLLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWELTILV 358
Db 299 TLLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQCGMTWVASWELTILV 358

```
Db 299 TILLVGRVAPARDKSKMVDPKKGTSLSDKLSFSHTHTLLGQFFQCGMGWVASWPLTILVL 358
Qy 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS 418
Qy 419 SYRVDLSLLGPKNFSGIIDLDLLELELQERLRLHQLVWSPAEARNISLODICVAPLNPD 478
Db 419 SYRVDLSLLGPKNFSGIIDLDLLELELQERLRLHQLVWSPAEARNISLODICVAPLNPD 478
Qy 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGTL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGTL 538
Qy 539 ALSMADYGAVPVFPFLAIGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 598
Db 539 ALSMADYGAVPVFPFLAIGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 598
Qy 599 EMRAFQRMAGMFOVTFTAERSLEDEINRTTAEDELPFATSVIVIFLYISLALGSYSWS 658
Db 599 EMRAFQRMAGMFOVTFTAERSLEDEINRTTAEDELPFATSVIVIFLYISLALGSYSWS 658
Qy 659 RVWVDSKATLGLGGVAVVLGAWMAAGFFSYLGRSSVLVQVVPFLVLSVGADNIFIV 718
Db 659 RVWVDSKATLGLGGVAVVLGAWMAAGFFSYLGRSSVLVQVVPFLVLSVGADNIFIV 718
Qy 719 LEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPMPAVRTFALTSG 778
Db 719 LEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPMPAVRTFALTSG 778
Qy 779 AVILDFLLQMSAFVALLSDSKROEASRLDVCCVKPQELPPGQEGGLLGGFFOKAYAP 838
Db 779 AVILDFLLQMSAFVALLSDSKROEASRLDVCCVKPQELPPGQEGGLLGGFFOKAYAP 838
Qy 839 FLLHWITRGW---LFLFLFGLVSLYMSCHISVGLDQELALPKDSYLLDYFLNRYF 894
Db 839 FLLHWITRGWVVPVPSQLLLFLALFVGLYMSCHISVGLDQELALPKDSYLLDYFLNRYF 898
Qy 895 EVGAPVYFVTTILGYNPFSSAGMNAICSSAGCNFSTQKIQVATEPPEOSYLAIPASSWV 954
Db 895 EVGAPVYFVTTILGYNPFSSAGMNAICSSAGCNFSTQKIQVATEPPEOSYLSLPASSWV 958
Qy 955 DFDIDWLTSSCCRLYISGNPKDKCPTVNSLNCCLKMSITMGSVRPSVSEQHKYLPW 1014
Db 959 DFDIDWLTSSCCRLYISGNPKDKCPTVNSLNCCLKMSITMGSVRPSVSEQHKYLPW 1018
Qy 1015 FLNDRNIIKCPKGLAAYSTSVNLTSDGOVLASRFMAHYKPKNSQDYTEALRAARELAA 1074
Db 1019 FLNDRNIIKCPKGLAAYSTSVNLTSDGOVLASRFMAHYKPKNSQDYTEALRAARELAA 1078
Qy 1075 NITADLRKVPGTDPAEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLD 1134
Db 1079 NITADLRKVPGTDPAEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCILLGLD 1138
Qy 1135 LRSGLLNLSISVIMLVDTVGFMAIWDI SYNNAVSLINLVAGMGSVEFVSHITRSPAISTK 1194
Db 1139 LRSGLLNLSISVIMLVDTVGFMAIWDI SYNNAVSLINLVAGMGSVEFVSHITRSPAISTK 1198
Qy 1195 PTLWERAKRATISMGSAVAGVANTMLPGILVLGLAKAQIQLIFFRLNLLITLGLLHG 1254
Db 1199 PTLWERAKRATISMGSAVAGVANTMLPGILVLGLAKAQIQLIFFRLNLLITLGLLHG 1258
Qy 1255 LVFLPVLISYVGPVDPVNPALAEQKRAEEA-----VAAMVVASCPNHPRSVST 1301
Db 1259 LVFLPVLISYVGPVDPVNPALAEQKRAEEGGGSHGGLAQITPPESPQLTTSMT 1313
```

RESULT 11

US-10-621-758A-2

; Sequence 2, Application US/10621758A

; Publication No. US20040093629A1

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

```
; APPLICANT: Wang, Luanan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-US
; CURRENT APPLICATION NUMBER: US/10/621.758A
; PRIORITY FILING DATE: 2003-07-17
; PRIORITY FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2
```

```
Query Match 78.5%; Score 5421.5; DB 4; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
```

```
Qy 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
Db 1 MAAAWL-GWLLWALLLSAAQGElyTPRHEAGVCTFYEECGKNPELSGSLTSLNVCSLSN 59
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTQ-ACCSAKQLVSLBASISITKALLTRCPCACSDN 119
Db 60 TPARKITGEHLLALQRICPRLYNGPNTTFFACSTYKQLSLLESSMSITKALLTRCPCACSDN 119
Qy 120 FVNHLCHNTCPNOSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAESYSCSRVRVPA 179
Db 120 FVSLHCHNTSPDQSLFINVTRVVERGAGEPPAVVAYEAFYQRSFAEKAYSCSVRIPA 179
Qy 180 AATLAVGTWCGYSGALCNQARWLNFQDGTGNGLAPLDITPHLLPBGAVGSGIOPNEG 239
Db 180 AASLAVGSMCGYSGALCNQARWLNFQDGTGNGLAPLDITPHLLPBGALPDGIQPLNGK 239
Qy 240 VARNCSOGDDVATCSQDCAASCPAIPARQALDSTFYLGQWPGSLVLIILICSVFVAVT 299
Db 240 IAPCNESQDDSVASCQDCAASCPVIPPPEALRPSFTWGRMPGWLALIIIFTAVFVLLS 299
Qy 300 ILLVGRVAPARDKSKMVDPKKGTSLSDKLSFSHTHTLLGQFPQGMGTWVASWPLTILVLS 359
Db 300 AVLVELRVVSNRNKNKAEGPQEA PKLPHKHLSPHTILGRFPQNGMGTWVASWPLTILVLS 359
Qy 360 VTPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS 419
Db 360 FTWVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRS 419
Qy 420 YRYDSLLLGPKNFSGILDLLELELQERLRLHQLVWSPAEARNISLODICVAPLNPD 479
Db 420 YKIDSLLLGSKNFSGILSLDFLELELQERLRLHQLVWSPAEARNISLODICVAPLNPN 479
Qy 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGTLA 539
Db 480 TSLDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFPLYCANAPLTFKDGTSLA 539
Qy 540 LSCMADYGAVPVFPFLAIGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 599
Db 540 LSCMADYGAVPVFPFLAIGYKGDYSEAEALIMTFSLNNTYPAGDPRLAQAKLWEEAFLE 599
Qy 600 MRAPORMAGMFOVTFTAERSLEDEINRTTAEDELPFATSVIVIFLYISLALGSYSWSR 659
Db 600 MESFORNTSDKFOVAFSAERSLEDEINRTIQDLPVFAVSYIIVFVLYISLALGSYSRCSR 659
Qy 660 VMVDSKATLGLGGVAVVLGAWMAAGFFSYLGRSSVLVQVVPFLVLSVGADNIFIVL 719
Db 660 VAVESKATLGLGVIIVLGAALAMGFYSYLGVPSSLVIIQVVPFLVLAAGDNIIFIVL 719
Qy 720 EYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGALTMPMPAVRTFALTSGLA 779
Db 720 EYQRLPRMFGQREAHIGRTLGSVAPSMMLCSLSEACFFLGALTMPMPAVRTFALTSGLA 779
```


1080	DLRKVP	GTDPNFEVFPYTTISNVFYQQYLTUTVPEGI	FTLALCFVTFVVCYLLGLDMCSG	1139	
1139	LNLLSI	VMILVDTVGFMA	LDISYNAVSLINLVS	AVGMSVEFVSHITRSFAISTKPTWL 1198	
1140	LNLLSI	IMILVDTIGLMA	WGISYNAVSLINLVT	AVGMSVEFVSHITRSFAVSTKPTRL 1199	
1199	ERAKAT	ISMGSAV	PAGVAMTNLP	PGILVLGLAKAQLIQIFFRNLNLLITLGLLHGLVFL 1258	
1200	ERAKAT	IVMGSAV	PAGVAMTNFP	CGILILGFAQAQLIQIFFRNLNLLITLGLLHGLVFL 1259	
1259	PVILSY	VGDPNPALAEQ	KRAEAAVAAVMV	ASCNHPHSRSTADNIYVNHSPFEGSI-KG 1317	
1260	PVILSY	LGDPVNQALVQ	EKLASEA-AVAPE	SCPOQPSPADAN--VNYGFAPELAHG 1316	
1318	AGAI	SNFLPNNGRQF	1332		
1317	ANAAR	SLPKSDQKF	1331		
RESULT 13					
US-10-646-301A-2					
; Sequence 2, Application US/10646301A					
; Publication No. US20040137467A1					
; GENERAL INFORMATION:					
; APPLICANT: Altmann, Scott W					
; APPLICANT: Wang, Luquan					
; APPLICANT: Graziano, Michael					
; APPLICANT: Murgolo, Nick					
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF					
; FILE REFERENCE: J01603-K1-US					
; CURRENT APPLICATION NUMBER: US/10/646,301A					
; CURRENT FILING DATE: 2003-08-22					
; PRIOR APPLICATION NUMBER: 60/397,442					
; PRIOR FILING DATE: 2002-07-19					
; PRIOR APPLICATION NUMBER: 10/621,758					
; PRIOR FILING DATE: 2003-07-17					
; NUMBER OF SEQ ID NOS: 50					
; SOFTWARE: PatentIn version 3.1					
; SEQ ID NO 2					
; LENGTH: 1331					
; TYPE: PRT					
; ORGANISM: Rattus sp.					
US-10-646-301A-2					
Query Match 78.5%; Score 5421.5; DB 4; Length 1331;					
Best Local Similarity 77.8%; Pred. No. 0;					
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;					
Qy	1	MAEAG	RGWLLWALLRLAQSE	PYTTTHQPGYCAFYDECGKNPELSGLMTLSNVCLSN 60	
Db	1	MAAAW	L-GWLLWALLLSAAQ	GElyTPKHEAGVCTFYEECGNPELSGGLTSLSNVCLSN 59	
Qy	61	TPARK	ITGDHLILLOKICP	RLYTCGNQ-ACCSAKOLVLEASISITKALLTRCPACSDN 119	
Db	60	TPARH	TGEHLALQRI	CPRLYNGPNTTFACSTKQLSLESMSITKALLTRCPACSDN 119	
Qy	120	FVNLC	HTCNSPNSQSLF	INTVTRVLAQLGAGQLPAVVAYEAFYQHSFABQSYDSCSRVRVPA 179	
Db	120	FVSLH	CHNTCSPDQSLF	INTVTRVVERGAGEPPAVVAYEAFYQRSFAEKAYESCQVRIPA 179	
Qy	180	AATLA	VGTCVYGSALCNAQRWLNFGQD	TGNGLAPLDTIFHLEPGQAVSGIOPLNEG 239	
Db	180	AASLA	VGSMCVYGSALCNAQRWLNFGQD	TGNGLAPLDTIFHLEPGQALPDGLOPLNGK 239	
Qy	240	VARN	ESGDDVATCSQDCAASCPATARP	ALDSTFVLGMPGSLVLIILCSVFAVVT 299	
Db	240	IAPCN	ESGDDSAVCSQDCAASCPIPP	PPALRPSFVWGRMPGWLALIIIFTAVFVLS 299	
Qy	300	ILLV	GFRVAPARDKSKMVDPKKGTSL	SDKLSFSHTLLGQFFQGGTGWASWPLTILVLS 359	
Db	300	AVLV	RLRVVNRNKNKAEGPQEA	PKLPHKHKLSPHITLGRFFQNWGTRVASWPLTVLALS 359	
Qy	360	VIPV	VALLAAGLVFTLTTD	PVELWSAPNSQARSEKAFHDQHF	GFPRNTQVILLTAPNRSS 419

RESULT 14
US-10-736-769-2
; Sequence 2, Application US/10736769
; Publication No. US20040161839A1
; GENERAL INFORMATION:

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRGMLLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
Db	3	AAWQGLLWALLRLNSAQGLYTHKAGCTFYEECGKNPELSGLTSLNSISCLNTPA	62
Qy	64	RKITGDHLLILQKICPRLYTGN-TQACCSAKQLVSLSEASLITKALLTRCPACSDNFVN	122
Db	63	RHVTGDHLLALQVCPRLYNGENDTYACCTKQLVSLDSSLITKALLTRCPACSENFVS	122
Qy	123	LHCHNTCSNQSFLFNVTVAQLGQOLPAVVAEAFYQHSFAQSDYSCSRVVPAAAT	182
Db	123	IHCNTCSPDOSFLFNVTVVQRPDQPLPAVVAEAFYQSPAKEYAESCVRVIPAAS	182
Qy	183	LAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPDITFHLLPEQQAADGKPLDGKITP	242
Db	183	LAVGSMCGVYGSALCNAQRWLNFGQDTGNGLAPDITFHLLPEQQAADGKPLDGKITP	242
Qy	963	P-SSCCRIYISGPNKPKCPSTVNSLNCIKNCMIITMGSVPSVEQHKYLPWFNDNPN	1021
Db	963	PSSSCCRIYTRGPHKDEFPCPSTDTSFNCLKNCMRTLFGVPRPTTEQHKYLPWFNDNPN	1022
Qy	1022	IKCPKGGIAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGGIAARTSVNLSSDQGIASQFMAHYKPLRNSQDFTALRASLLAANITAEIR	1082

Qy	1082	KVPGTDPAPEVPPTTITNVPYEQVLTILPEGLFMLSCLVPTFAVSCLLIGLDRSGLN	1141
Db	1083	KVPGTDPEFVFPPTTISNVFYQQLTLVLPBGITFLALCFVPTFVVCVLLIGLDRSGLN	1142
Qy	1142	LLSIVMLVDTVGFMALWDISYNVSLINIVSAGMSVEFVSHITRSFAISTKPTWLEA	1201
Db	1143	LLSIIIMLVDTIGLMAVMSISYNVSLINLVAGMSVEFVSHITRSFAVSTKPTRLERA	1202
Qy	1202	KEATISMGSAVFAVAVMTNLPGLIVLGLAKAQILQIFFFRLNLLITLGLLHGLVFLPVI	1261
Db	1203	KDATIFMGSAVFAVAVMTNFPGLIILGFAQLIQLIFFRLNLLITLGLLHGLVFLPVI	1262
Qy	1262	LSYVGPDPNPALEQKRAEAAVAVNVASCPNHPRSVSTADNINYNHSEGS-IKGAGA	1320
Db	1263	LSYVGPDPNPALEQKRAEAAVAVNVASCPNHPRSVSTADNINYNHSEGS-IKGAGA	1320
Qy	1321	ISNPLPNNGRQF 1332	
Db	1322	ASSSLPKSDQKF 1333	
RESULT 16			
US-10-663-208A-12			
; Sequence 12, Application US/10663208A			
; Publication No. US20040132058A1			
; GENERAL INFORMATION: Scott W			
; APPLICANT: Altmann, Scott W			
; APPLICANT: Wang, Luquan			
; APPLICANT: Graziano, Michael			
; APPLICANT: Murgolo, Nick			
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF			
; FILE REFERENCE: JB01603K2 US			
; CURRENT APPLICATION NUMBER: US/10/663,208A			
; PRIORITY FILING DATE: 2003-09-16			
; PRIOR APPLICATION NUMBER: 60/397,442			
; PRIOR FILING DATE: 2002-07-19			
; PRIOR APPLICATION NUMBER: 10/621,758			
; PRIOR FILING DATE: 2003-07-17			
; PRIOR APPLICATION NUMBER: 10/646,301			
; PRIOR FILING DATE: 2003-08-22			
; NUMBER OF SEQ ID NOS: 50			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 12			
; LENGTH: 1333			
; TYPE: PR1			
; ORGANISM: Mus sp.			
US-10-663-208A-12			
Query Match 78.3%; Score 5407; DB 4; Length 1333;			
Best Local Similarity 77.2%; Pred. No. 0;			
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;			
Qy	4	AGLRGMLLWALLRLAQSEPYTHIOPGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA	63
Db	3	AAWQGLLWALLRLNSAQGLYTHKAGCTFYEECGKNPELSGLTSLNSISCLNTPA	62
Qy	64	RKITGDHLLILQKICPRLYTGN-TQACCSAKQLVSLSEASLITKALLTRCPACSDNFVN	122
Db	63	RHVTGDHLLALQVCPRLYNGENDTYACCTKQLVSLDSSLITKALLTRCPACSENFVS	122
Qy	123	LHCHNTCSNQSFLFNVTVAQLGQOLPAVVAEAFYQHSFAQSDYSCSRVVPAAAT	182
Db	123	IHCNTCSPDOSFLFNVTVVQRPDQPLPAVVAEAFYQSPAKEYAESCVRVIPAAS	182
Qy	183	LAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPDITFHLLPEQQAADGKPLDGKITP	242
Db	183	LAVGSMCGVYGSALCNAQRWLNFGQDTGNGLAPDITFHLLPEQQAADGKPLDGKITP	242
Qy	243	CNCSQGGDDVATCSQDCCAASCPAARQALDSTFYLGQMPGSLVLIILCSVFAVVTILL	302
Db	243	CNCSQGGDSASCSQDCCAASCPVIPPALRPSFYMGMRPGWLALIIIFTAVFVLLSVL	302

```
QY 303 VGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQPFQCGWGTWVASHWPLTILVLVSVIP 362
Db 303 VYLRVASNRNKNKTAGSQEAPNLPKRPRFSPHTVLGRPFESWGTVRVASHWPLTVLALSFI 362
QY 363 VVALAAGLVFTLTDPVELMSAPNSOARSEKAFHQHFGPFRPTNQVILTAAPNRSSVY 422
Db 363 VIALSVGLTFIELTDPVELMSAPKQARKEKAFHDEHFGPFRPTNQIPVTAKNRSSYK 422
QY 423 DSSLGPKNPSGILDLDLLELLELQERLRLHQLVMSPEAQNRNLSLODICVAPLNPNTSL 482
Db 423 DSSLGPKNPSGILSDLLQELLELQERLRLHQLVMSHEAQNRNLSLODICVAPLNPNTSL 482
QY 483 YDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKDGTAALASC 542
Db 483 TDCCVNSLLQYFQNNHTLLLTANQTLNGQTSQVDWKDHLFYCANAPLTYKDGTAALASC 542
QY 543 MADYGAPVPFPLAIGYKGDYSEAEALJMTFSLNNYPAGDPLAQAKLWEEAFLEEMRA 602
Db 543 IADYGAPVPFPLAVGYQGTDYSEAEALITFSINNPADDPMAHAKLWEEAFLEEMQS 602
QY 603 FORMMAGMPQVTTFAERSLEDEINRTTAEDLPFIPATSYIIVILYISLALGSYSSSRVMV 662
Db 603 FORSTADKFOIAPSAERSLEDEINRTTIQDLVPFAISYILVFLYISLALGSYSRWSRVAV 662
QY 663 DSKATILGCGVAVVLGAVMAAGFFSYLGRSLVILQVVPFHLVLSVGADNIPFVLEYO 722
Db 663 DSKATILGCGVAVVLGAVVAAGFFSYLGVPSLVIIQVVPFHLVLAAGADNIPFVLEYO 722
QY 723 RLPRRPGEBREHIGALGRVAPSMLLCSLSAICFFLALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMEGOREAHIGTLGSAVAPSMLLCSLSAICFFLALTPMPAVRTFALTSGLAIVL 782
QY 783 DFLQMSAFVALLSDSKRQESRLDVCCKVPQELPPGQEGGLLGFQKAYAPFLIH 842
Db 783 DFLQMTAFVALLSDSKQESRPDVCCFSSRNLPPEKQEGGLLGFQKAYAPFLIH 842
QY 843 WITRGVLLFLFALFGVSLYSWCHISVGLDQELAPKDSYLLDYFLFLNRYEFGAPVYF 902
Db 843 RPIRPVLLFLFALFGANILCMNISVGLDQELAPKDSYLLDYFLFLNRYLEVGPVYF 902
QY 903 VTTLGVNFSSEAGMAICSSAGCNFSTOKIYATERPEQSVLAIAPASSWDDFDIWL 962
Db 903 DTSGVNFSTEAGMAICSSAGCESFSLTKIYASBFPNQSYVAIAASSWDDFDIWL 962
QY 963 P-SSCCRLYISGPNKDFCPSVTNLSNCLKNQMSITMGSVRPSVEQFHKYLFPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDFCPSDTTSNCLKNQMSITMGSVRPSVEQFHKYLFPWFLNDRN 1022
QY 1022 IKCPKGLLAAYTSVNLSDGQVLAASFMAHKLKNSQDYTEALRAARELANITADLR 1081
Db 1023 IRCPKGLLAAYTSVNLSDGQIIASQFMAHKLKNSQDYTEALRAARELANITADLR 1082
QY 1082 KYPGTDPAPEVPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLIGLIDLSGLIN 1141
Db 1083 KYPGTDPAPEVPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLIGLIDLSGLIN 1142
QY 1142 LLSIIMILVDITVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAIKTPWLER 1201
Db 1143 LLSIIMILVDITVGLMAVWGISYNAVSLINLVSAGMSVEFVSHITRSPAIKTPWLER 1202
QY 1202 KEATTSMGSAVPAGVAMTNLPGLVGLAKAQLIQIPFRLNLLITLGLHGLVFLPV 1261
Db 1203 KDATTSMGSAVPAGVAMTNFPGLVGLAKAQLIQIPFRLNLLITLGLHGLVFLPV 1262
QY 1262 LSVYGPDPNPALALQKRAEBAVAVMVASCPNHFSRVSTADNIVNHSFEGS-1KGAGA 1320
Db 1263 LSVYGPDPNPQALVLEKALATEA-AMVSEPCQYPPPADANTSIVNVPFEPFEPINA 1321
QY 1321 ISNFLPNNGROF 1332
Db 1322 ASSSLPKSDQKF 1333
```

```
RESULT 17
US-10-646-301A-12
; Sequence 12, Application US/10646301A
; Publication No. US20040137467A1
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-646-301A-12
```

Query Match 78.3%; Score 5407; DB 4; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

```
QY 4 AGLRGWLLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCISLNTPA 63
Db 3 AAWQGWLLWALLRLNLAQAGELYPTHKAGCTFYECCKNPELSGGLTSLNSICLSLNTPA 62
QY 64 RKITGDHLILQKICPLRYTGN-TCACCSAKQLVLEASLSTTKALLTRCPACSNFVN 122
Db 63 RHVTGDHLALQRCVCPRLYNGPNDTYACCSTKQLVSLDSSLSTTKALLTRCPACSNFVS 122
QY 123 LHCHNTCSPNQSLFTNVTTRVAQAGQPAVVAEAFYQHSFAEQSYDSCSRVRVPAAT 182
Db 123 IHCHNTCSPDQSLFTNVTTRVQDQGLPAVVAEAFYQHSFAEQSYDSCSRVRVPAAS 182
QY 183 LAVGTMCVGYGSAALCNAQRLNFGQDGTGNGLAPLITFHLBPQAVGSGIQLPNEGVAR 242
Db 183 LAVGSMCVGYGSAALCNAQRLNFGQDGTGNGLAPLITFHLBPQAVGSGIQLPNEGVAR 242
QY 243 CNESGDDVATCSQDCAASCPAIAEPQALDSTFVYGQMPGSLVLIILCSVPAVVTILL 302
Db 243 CNESGQSDSAACSCQDCAASCPVIPPALRPSFTYGRMPGWLIIIFTVFVLLSVVL 302
QY 303 VGRVAPARDKSKWDPKKGTSLSDKLSPSTHTLLGQPFQCGWGTWVASHWPLTILVLVSVIP 362
Db 303 VYLRVASNRNKNKTAGSQEAPNLPKRPRFSPHTVLGRPFESWGTVRVASHWPLTVLALSFI 362
QY 363 VVALAAGLVFTLTDPVELMSAPNSOARSEKAFHQHFGPFRPTNQVILTAAPNRSSVY 422
Db 363 VIALSVGLTFIELTDPVELMSAPKQARKEKAFHDEHFGPFRPTNQIPVTAKNRSSYK 422
QY 423 DSSLGPKNPSGILDLDLLELLELQERLRLHQLVMSPEAQNRNLSLODICVAPLNPNTSL 482
Db 423 DSSLGPKNPSGILSDLLQELLELQERLRLHQLVMSHEAQNRNLSLODICVAPLNPNTSL 482
QY 483 YDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKDGTAALASC 542
Db 483 TDCCVNSLLQYFQNNHTLLLTANQTLNGQTSQVDWKDHLFYCANAPLTYKDGTAALASC 542
QY 543 MADYGAPVPFPLAIGYKGDYSEAEALJMTFSLNNYPAGDPLAQAKLWEEAFLEEMRA 602
Db 543 IADYGAPVPFPLAVGYQGTDYSEAEALITFSINNPADDPMAHAKLWEEAFLEEMQS 602
QY 603 FORMMAGMPQVTTFAERSLEDEINRTTAEDLPFIPATSYIIVILYISLALGSYSSSRVMV 662
Db 603 FORSTADKFOIAPSAERSLEDEINRTTIQDLVPFAISYILVFLYISLALGSYSRWSRVAV 662
```

```
QY 663 DSKATLGLGGVAVVLGAVMAAMGFPSYIGIRSSILVILQVVPFLVLSVGCADNIFIFVLEYQ 722
Db 663 DSKATLGLGGVAVVLGAVMAAMGFPSYIGVPSLSLVIQVVPFLVAVGADNIFIFVLEYQ 722
QY 723 RLPRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGLALTMPMPAVRTFALTSGLAVAL 782
Db 723 RLPRMPGQREAHIGRTLGSVAPSMMLCSLSEAICFFLGLALTSMMPAVRTFALTSGLAIF 782
QY 783 DFLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGLLGLFFQKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQREASRLDVCCVSSRNLPKPQKQEGLLGLFFQKAYAPFLH 842
QY 843 WITRGVLLFLALFGVSLYSWMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPVYF 902
Db 843 RTIRPVLLFLFLFANLYLMCNISVGLDQDALPKDSYLLDYFLFLNRYLEVGPVYF 902
QY 903 VTTLGVNFSSEAGMNAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGNFSSEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT 962
QY 963 P-SSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRPSVEQPHKYLPMFLNDRPN 1021
Db 963 PSSCCRIYTRGPHKDEFCPTDTSFNCLKNCWNRITLGVPRPTTBOFHKYLPMFLNDRPN 1022
QY 1022 IKCPKGLAAYSTVNLTSDGVLASRPMAYHKPLKNSODYTEALRAARELAANITADLR 1081
Db 1023 IRCPKGLAAYSTVNLTSDGQIIASQFMAYHKPLRNSQDTEALRASRLAANITAEUR 1082
QY 1082 KVPGETDPAVEVPYITTVFYEQYLTILPEGLPMLSLCLVPTFAVSCLLGLDLRSGLIN 1141
Db 1083 KVPGETDPEVEVPYITTVFYEQYLTILPEGLFTLALCFVTFVVCYLLGLDIRSGLIN 1142
QY 1142 LLSIVMLVDYTFGMALWDISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLER 1201
Db 1143 LLSIILVDYITGLMAVWGISYNAVSLINLTVAGMSVEFVSHITRSFAVSTKPTRLERA 1202
QY 1202 KENTISMGAVPAGVANTWLPGLVTLGLAKAQLIQIFFERLNLITLLGLLHGLVPLVI 1261
Db 1203 KOATIFMGSAVAGVANTWFPGLVTLGLGFAQAQLIQIFFERLNLITLLGLLHGLVPLV 1262
QY 1262 LSVGVDPNPALALEOKRAEEAAVAVMVASCPNHPSRVSTADNIYVNHSEFSGS - IKG 1320
Db 1263 LSVGLGPDVQNALVLEKLATEA-AMVSEPCQYPPPADANTSDYVNGFNPFEIPEINA 1321
QY 1321 ISNFLPNNGRQF 1332
Db 1322 ASSSLPKSDQKF 1333
```

RESULT 18

```
US-10-736-769-12
; Sequence 12, Application US/10736769
; Publication No. US20040161838A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Luquan
; APPLICANT: Altmann, Scott W
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
```

```
; LENGTH: 1333
; TYPE: PR1
; ORGANISM: Mus sp.
US-10-736-769-12
```

```
Query Match 78.3%; Score 5407; DB 4; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;
```

```
QY 4 AGURGLWLLALLLAQSEBPTTHHQGYCAPFYDECKNPELSGLMTLSNVSCLSNTPA 63
Db 3 AAMQGLWLLALLLSAQSELYTPTHKAGFCFTFBECKNPELSGGLTSLNSISCLSNTPA 62
QY 64 RKITGDHLILLOKICPLRYLTGPN--TOACCSAKOLVLSLEASISITKALITRCPACSDNFVN 122
Db 63 RHVTGDHLLALQRCVPRLYNGPNDTYACCSTKQLVSLDSSLISITKALLTRCPACSENFVS 122
QY 123 LHCHTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFASQSDYSCSRVRVPAAT 182
Db 123 IHCHTCSPDQSLFINVTRVVRDPGQLPAVVAEAFYQHSFASQSDYSCSRVRIPAAAS 182
QY 183 LAVGTCGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLLPEGOAVGSGIQPLNEGVAR 242
Db 183 LAVGSMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITPHLLPEGOALADGKPLDKTTP 242
QY 243 CNESQGDVATCSQDCAASCAPAIARPOALDSTFYLGQMPGSLVLIILILCSVFAVVITLL 302
Db 243 CNESQGEDSAACSQDCAASCVPVPPPALRPSFVMGRMPCWALIIIFTAVFVLLSVVL 302
QY 303 VGRVAPARDKSMVDPKKGTSLSDKJSFSTHTLLGQFQGWGTWVASWPLTILVLSVIP 362
Db 303 VVLRVASNRKNKTAGSQEAPNLPRKRFRSPHTVLGRFESWGTRVASWPLTVLALSIV 362
QY 363 VVALAAGLVFTELTTDPVELASAPNSQARSKAFHQHFGFPFRTNQVILTAPNRSSVRY 422
Db 363 VIALSVGTLFTELTTDPVELWSAPKSQARKKAFHDEHFGFPFRNQVILFVIAKNRSSYKY 422
QY 423 DSLILGPKNFGSLDLILLLELLEQLERLRLHQLVMSPEAQRNISLQDICYAPLNPDTSL 482
Db 423 DSLILGPKNFGSILDLILLLELLEQLERLRLHQLVMSPEAQRNISLQDICYAPLNPHTSL 482
QY 483 YDCCINSLLQVFNQNRITLTLTANQTLMGQTSQVQDKHFLYCANAPLTKDGTALALSC 542
Db 483 TDCCVNSLLQVFQNNHTLTLTANQTLNGQTSQVQDKHFLYCANAPLTKDGTALALSC 542
QY 543 MADYCAPVPPFLAIGYKGDYSEAEALIMTFLSNYPAGDPLRAQAKLWEEAFLEEMRA 602
Db 543 IADYCAPVPPFLAVGGYQGTDIYSEAEALITFESINNYPADDDPRMAHAKLWEEAFLEEMQS 602
QY 603 FORRMAGMFQVTTAERSLEDEINRTTADLPFIATSYIVIFYITSLALGSYSSSRVWV 662
Db 603 FORSTADKQFIAPSAERSLEDEINRTTIQDLVPFAISYLVIFYITSLALGSYSSSRVAV 662
QY 663 DSKATLGLGGVAVVLGAVMAAMGFPSYIGIRSSILVILQVVPFLVLSVGCADNIFIFVLEYQ 722
Db 663 DSKATLGLGGVAVVLGAVMAAMGFYSYLGVPSSLVIQVVPFLVAVGADNIFIFVLEYQ 722
QY 723 RLPRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGLALTMPMPAVRTFALTSGLAVAL 782
Db 723 RLPRMPGQREAHIGRTLGSVAPSMMLCSLSEAICFFLGLALTSMMPAVRTFALTSGLAIF 782
QY 783 DFLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGLLGLFFQKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQREASRLDVCCVSSRNLPKPQKQEGLLGLFFQKAYAPFLH 842
QY 843 WITRGVLLFLALFGVSLYSWMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPVYF 902
Db 843 RTIRPVLLFLFLFANLYLMCNISVGLDQDALPKDSYLLDYFLFLNRYLEVGPVYF 902
QY 903 VTTLGVNFSSEAGMNAICSSAGCNFSTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGNFSSEAGMNAICSSAGCESFSLTQKIQYASEFPNQSYVAIAASSWVDDFIDWLT 962
```



```
QY 963 P-SSCCRLYISGNKDKFCBPTVNSLNCILKNCMSITMGSVRPSVEQPHKYLPHFLNDREN 1021
Db 963 PSSCCRIYTRGPHKDFECPSTDTSFNCLKNCMNRTILGVPVPTTEQPHKYLPHFLNDTEN 1022
QY 1022 IKCPKGGLAAYSTVNLTSQGVLAGRFMAYHKPLKNSODYTEALRAARELAANITADLR 1081
Db 1023 IRCPKGLAAYRTSVNLSSDQIIASQFMAYHKPLNSQDFTALRASRLAANITABLR 1082
QY 1082 KVPGTDAPEVFPYTTITNVPEQYLTILPEGLFMLSCLVPTPAVSCLLGLDLRSLN 1141
Db 1083 KVGTDNPEVFPYTTISNVFYQYLTALPEGITLALCFVPTVVCVLLGLDIRSGILN 1142
QY 1142 LLSIVMLVDTYVCFMADWISYNAVSLINLVSAGVMSVEFVSHITSPALSTKPTWLER 1201
Db 1143 LLSIVMLVDTYVCFMADWISYNAVSLINLVSAGVMSVEFVSHITSPALSTKPTWLER 1202
QY 1202 KBATISMGSAVAGVANTNLPGILVLGLAKAQLIQIFPRLNLLITLLGLLHGLVLPVI 1261
Db 1203 KDATIFWGSNAVAGVANTWFPGLIILGPAQAQLIQIFPRLNLLITLLGLLHGLVLPVI 1262
QY 1262 LSYVGPDPNPALALEQKRAEBAVAAMVNASCPNHPRSVSTADNIYVNHSPFEGS- 1320
Db 1263 LSYVGPDPNPALALEQKRAEBAVAAMVNASCPNHPRSVSTADNIYVNHSPFEGS- 1321
QY 1321 ISNFLPNNGRQF 1332
Db 1322 ASSSLPKSQKF 1333

RESULT 19
US-10-450-763-53050
; Sequence 53050, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53050
; LENGTH: 982
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(982)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53050

Query Match 64.6%; Score 4466; DB 5; Length 982;
Best Local Similarity 98.0%; Pred. No. 0;
Matches 867; Conservative 1; Mismatches 9; Indels 8; Gaps 3;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDYSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDYSCSRVRPAA 180
```

```
QY 181 ATLAVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
Db 181 ATLAVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
QY 241 ARCNSQGDVATCSCODCAASCP--AIARQALDSTFFVLGQMPGSLVLIILICSPAVV 298
Db 241 ARCNSQGDVATCSCODCAASCP--AIARQALDSTFFVLGQMPGSLVLIILICSPAVV 298
QY 299 TILLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQPFQGGTGWVWASPLTILVL 358
Db 299 TILLVGRFVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQPFQGGTGWVWASPLTILVL 358
QY 359 SVIPVVALAAGLVFTTELTTPVELMSAPNSQARSEKAFHQHFGPPFRFNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTELTTPVELMSAPNSQARSEKAFHQHFGPPFRFNQVILTAPNRS 418
QY 419 SYRYSLLGPKNFSGILDLLELLELOERLHLQVMSPEAQRNLSLQDICYAPLNDP 478
Db 419 SYRYSLLGPKNFSGILDLLELLELOERLHLQVMSPEAQRNLSLQDICYAPLNDP 478
QY 479 NTSLYDCCINSLLQVFNQNRNRTLLLTANOTLMGQTSQVDMKDFLYCANAPLTFKDGTL 538
Db 479 NTSLYDCCINSLLQVFNQNRNRTLLLTANOTLMGQTSQVDMKDFLYCANAPLTFKDGTL 538
QY 539 ALSCMADYGAFFPFLAIGGYKGDYSEAEALIMTFLSNYPAGDPRLAQAKLWBEAPLE 598
Db 539 ALSCMADYGAFFPFLAIGGYKGDYSEAEALIMTFLSNYPAGDPRLAQAKLWBEAPLE 598
QY 599 ENRAFQRMAGMFQVTFPAERSLEDEINRTTAEDLPFATSYVIVIPLYTSLAGSYSSWS 658
Db 599 ENRAFQRMAGMFQVTFPAERSLEDEINRTTAEDLPFATSYVIVIPLYTSLAGSYSSWS 658
QY 659 RVWVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSVLQVWPVLVSGADNIFIFV 718
Db 659 RVWVDSKATLGLGGVAVVLGAVMAAGPFSYLGIRSSVLQVWPVLVSGADNIFIFV 718
QY 719 LEYQRLRRRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGALTPMPAVRTFALTSG 778
Db 719 LEYQRLRRRPGPREVHIGRALGRVAPSMMLCSLSEAICFFLGALTPMPAVRTFALTSG 778
QY 779 AVILDFLOMSAFVALLSDSKROBASRLDVCCCKVQBELPPPGQEGILLGFPQKAYAP 838
Db 779 AVILDFLOMSAFVALLSDSKROBASRLDVCCCKVQBELPPPGQEGILLGFPQKAYAP 838
QY 839 PLLHWITRGW---LLLFLALFGVSLYSYMSCHISVGLDQELALPK 879
Db 839 PLLHWITRGWVPSQLLLFLALFGVSLYSYMSCHISVGLDQELALPK 883

RESULT 20
US-10-208-731-2
; Sequence 2, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-208-731-2
```

[illegible][illegible]

```
Db 385 SSOARLEKEYFDHGFQGFRTQIIRAPLTKHIIYQIPSPGADVPFGPDLQIILHQLV 444
Qy 446 ELQERLRHLQVMSPEARNIQLQDICYAPLNDNTSLYDCCINSILQYFQNNRTLLILTA 505
Db 445 DLQIAIEN--ITASVDNETVTLLQDICALPLSPYNT---NCTILSVLYNFQNSHVLHDKK 499
Qy 506 NQTLMGQTSQVDKDHFLYCANAPLTPKDGITALALSCHADYCAPVPPFLAIGYKGYOYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLIHDCLGTGCGPVPFVPLVGLGYDDQNYN 556
Qy 566 EBAELIMTSLNYPAGDPRLAQAKLWEAFLEEMRAFQRRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFINFVKNYK---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVMDVKATLGLGVAIVLGVMAAAG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHIKSCRLLVDSKVSIGIAGIILVLSVACSIG 673
Qy 686 FSYGLIRSSLVILQVVPFLVSVGADNIFIVLEYQRLPRPGEPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQYQDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSAICFFLGLATPMPAVRTFALTSLGLAVILDLELLQMSAFVALLSLDSKQRAS 805
Db 734 SMLFSSFTSEVAFPLGALSVMFAVHTFSLFAGLAVFDLFLQITCFVSLGLGDIKQKKN 793
Qy 806 RLDVCCVKPQBLPPPGQ--GEGILLGFFQKAYAPFLHMTIRGVLLFLALFVGLSYSM 864
Db 794 RLDIFCCVGAEDGTSVQASSECLFRFFKNSYSPILLKDWMPRIVIAFVGVLSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFNRYFVGAPVYFVTTLGYNFSSBAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLERGHDYTSKQGMVCGMG 913
Qy 925 CNMFSEFQKIQVATEPEQSYLAIAPASSWDDFDLWLP--SSCRLYISGPNKDFECPST 983
Db 914 CNDSLVQOIFNAQLDNTRYIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKMCSIT--MGSVRPSVEQFKYLPFLNDRPNKCPKGLAAYSTSVN--LTS 1040
Db 971 VDPACVR--CRPLTEGKQRPQGGDMRFLPFLSDNPNPKCGKGHAAYSAAVAILLGH 1029
Qy 1041 DQGVLASRPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
Db 1030 GTRVGATFMTYHTVLOTSAFDALKKARLIASNVT--ETWINGS--AYRVEPYSVFV 1086
Qy 1101 FYEQYLITLPEGLFMLSCLVTFPAVSCLLGLDLSGLNLLSIVMLIVDTVGPMALWD 1160
Db 1087 FYEQYLLTIDDTIFNLGVSLSAIFLVTVMVLLGCELWSAVIMCATTIAMVLNFMGMWLMG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRGSFAISTKPTMLERAKKATISMGSAVPAVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVEREAEALAHMGSSVFSGITLK 1206
Qy 1221 LRGILVLGLAKAQIQLIFFRNLNLITLGLHLGLVFLPVLSYVGPDPNPA 1272
Db 1207 FGGIVVLAPAKSIQIFIFVFRMYLAWMLLGLATHGLIFLPLVLSYIGSPSNKA 1258

RESULT 22
US-10-741-600-1542
; Sequence 1542, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
```

```
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1542
```

```
Query Match 34.8%; Score 2402.5; DB 5; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.2e-199;
Matches 522; Conservative 237; Mismatch 450; Indels 103; Gaps 25;
```

```
Qy 7 RGMILWALLLRLLAQSEPTTTHQPGVCAPFYDRSG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLAAGLGLLLLCAPQFSQ-----SCWYGGCGIAYGDKRYNCBSYG-----46
Qy 60 NTPARKITGDHLILQIKCPRLYTGTQACCSAKQLVSLREASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFG--NVLSCDQVQLQTLKDLNLQLPLQFLSRCSCFPN 103
Qy 120 FVNLCHNTCSNQSLFINVTR-----VAQLGAGQLPAAVAYEAFYQHSFABOSYSCSRV 175
Db 104 LILNLCFELTCSPROQFLNVTATEDYVDPVNTQTKTNKELQYVVGQSPANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCVYGSALCNAQRWLNFGQDTGNGLAPLDIT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNWIYEMFNKONGQAPFTTTFVFSDFPVH-----215
Qy 229 VSGIQPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNNATKGDSEVDVATPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
Qy 280 QMPSGLVLIILCSVPFVAVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
Db 273 WITTMAPLLVFGAFVAVCKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
Qy 327 DKLSESTHILQPPQGHGTWASWPLILVSLVPPVVALAAGLVTELTTPDELWSAP 386
Db 325 DPVSAAFEGCLRLRFLTRGSCFVRNPGCVIPFSLVFITACSSGLFVFRVTTNPVLDWSAP 384
Qy 387 NSQARSEKAFIDHQHGFPPFTNQVLTITAPNRSSRYDSLLLGPK--NFSGILDLLELL 445
Db 385 SSOARLEKEYFDHGFQGFRTQIIRAPLTKHIIYQIPSPGADVPFGPDLQIILHQLV 444
Qy 446 ELQERLRHLQVMSPEARNIQLQDICYAPLNDNTSLYDCCINSILQYFQNNRTLLILTA 505
Db 445 DLQIAIEN--ITASVDNETVTLLQDICALPLSPYNT---NCTILSVLYNFQNSHVLHDKK 499
Qy 506 NQTLMGQTSQVDKDHFLYCANAPLTPKDGITALALSCHADYCAPVPPFLAIGYKGYOYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLIHDCLGTGCGPVPFVPLVGLGYDDQNYN 556
Qy 566 EBAELIMTSLNYPAGDPRLAQAKLWEAFLEEMRAFQRRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEKFINFVKNYK---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPFATSYIVIFLYISLALGSYSMSRVMDVKATLGLGVAIVLGVMAAAG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHIKSCRLLVDSKVSIGIAGIILVLSVACSIG 673
Qy 686 FSYGLIRSSLVILQVVPFLVSVGADNIFIVLEYQRLPRPGEPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQYQDERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSAICFFLGLATPMPAVRTFALTSLGLAVILDLELLQMSAFVALLSLDSKQRAS 805
Db 734 SMLFSSFTSEVAFPLGALSVMFAVHTFSLFAGLAVFDLFLQITCFVSLGLGDIKQKKN 793
Qy 806 RLDVCCVKPQBLPPPGQ--GEGILLGFFQKAYAPFLHMTIRGVLLFLALFVGLSYSM 864
Db 794 RLDIFCCVGAEDGTSVQASSECLFRFFKNSYSPILLKDWMPRIVIAFVGVLSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFNRYFVGAPVYFVTTLGYNFSSBAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLERGHDYTSKQGMVCGMG 913
```

QY 925 C N N F S F T Q K I Q V A T E F F P Q S Y L A I P A S S W V D D F I D W L T P - S S C R L Y I S G P N K D K F C P S T 983
D B 914 C N N D S L V Q Q I F N A A Q L D N Y T R I G F A P S S W I D D Y F D W K P Q S S C R V - - - D N I T D Q F C N A S 970
QY 984 V N S N C L K N C M S I T - M G S V R P S V E Q F H Y L P W F L N D R P N I K C P K G L A A Y S T S V N - - L T S 1040
D B 971 V V D P A C V R - C R P L T P E G K Q P Q G D F M R F L P M F L S D N P N P K C G K G H A A Y S S A V N I L L G H 1029
QY 1041 D G Q V L A S R F M A Y H K P L K N S Q D Y T E A L R A A R E A L A N I T A D L R K V P G T D P A F E V F P Y T I T N V 1100
D B 1030 G T R V G A T Y F M T Y H T V L Q T S A D F I D A L K K A R L I A S N V T - E T W G I N G S - - A Y R V P P Y S V F Y V 1086
QY 1101 F Y E Q Y L T I P E G L F M L S C L V P T F A V S C L L G L D L R S G L L N L L S I V M L V D T V G F M A L W D 1160
D B 1087 F Y E Q Y L T I I D D T I F N L G S L G A I F L V T W V L L G C E L W S A V I M C A T I A M V L N M F G V M W L M G 1146
QY 1161 I S Y N A V S L I N L V S A V G M S V E F V S H I T R S F A I S T K P T W L E R A K E A T I S M G S A V E A G V A M T N 1220
D B 1147 I S L N A V S L V N L V M S C G I S V E F C S H I T R A F T V S M K G S R V E A E A L A H M G S S V F S G I T L T K 1206
QY 1221 L P G I L V L G A K A Q L I Q I F F R L N L L I T L L G L H G L V F L P V I L S Y V G P D V N P A 1272
D B 1207 F G G I V V L A F A K S Q I F Q I F F R M Y L A W L L G A T H G L I F L P V L L S Y I G P S V N K A 1258
RESULT 24
US-10-756-149-4924
; Sequence 4924, Application US/10756149
; Publication No. US20050181375A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: FILE
; CURRENT APPLICATION NUMBER: US/10/756,149
; CURRENT FILING DATE: 2004-01-12
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4924
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4924
Query Match 34.8%; Score 2402.5; DB 5; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.2e-199;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
QY 7 R G W L L A L L R L A Q S E P Y T T I H Q P G Y C A F Y D E C G - - - - - K N P L S G S L M T L S N V S C L S 59
D B 4 R G L A L G L L L L L C P A Q V F S Q - - - - - S C V W Y G E C G I A V G D K R Y N C E Y S G - - - - - 46
QY 60 N T P A R K I T G D H L I L L Q K I C P R L Y T G P N Q A C C S A K Q L V S L E A S I S I T K A L L T R C P A C S D N 119
D B 47 - - P P K P L P K D G Y D L V Q S L C P G F F G - N V S L C C D V R Q L Q T L K O N L Q L P L Q S L R C P S C F Y N 103
QY 120 F V N L H C H T C S P N Q S L F I N T R - - - - - V A Q L G A Q L P A V V A Y E A F Y O H S F A E Q S Y D S C S R V 175
D B 104 L L N L F C E L T C S P R O S Q L N V T A T E D Y D P V T N Q T K T N K E L Q Y V G S F A N A M Y N A C R D V 163
QY 176 R V P A A A T L A V G T M C G V Y G S A L C N A Q R W L N F Q G D T G N G L A P L D I T - - - - - F H L L E P G Q A 228
D B 164 E A P S N D K A L G L L C K D A D A - C N A T N M I E Y M F N K D N G O A P P T I T P V P S D P P V H - - - - - 215
QY 229 V G S G I Q P L E G V A R C N S Q G D D V A T C Q D C A A S C P A I A R P Q - - - - - A L D S T F Y L G 279
D B 216 - - - G M E P M N A T K G C D S V D E V T A P C S Q D C S I V C G P K P Q P P P A P W T I L G L D A M Y I M 272
QY 280 Q M P G S L V L I I L C S V E A V V T I L L - - - - - V G F R V A P A R D K S K M V D P K K G T S L S 326
D B 273 W I T Y M A F L L V F F G A F A V M C Y R K R Y F V S E Y T P I D S N I A F S V - N A S D R G E - - - - - A S C C 324

QY 327 D K L S P S T H T L L G Q F O G W G T W A S W P L T I L V S I P V V A L A A G L V F T E L T T D P V E L M S A P 386
D B 325 D P V S A A P E G C L R L F T R W S F C V R N P C C V I F F S L V I T A C S S G L V F V R V T T N P V D L M S A P 384
QY 387 N S Q A R S E K A P H D Q H G P F F R T N Q V I L T A P N R S S Y R Y D S L L L G P K - N P S G I L D L L L E L L 445
D B 385 S S Q A R L E K E Y F D Q H G P F F R T E Q L I I R A P L T D K H I T Q P Y P S G A D V P F G P L D I Q I L H O V L 444
QY 446 E L Q E R L H L Q W S P E A Q R N I S L O D I C V A P L N P O N T S L Y D C C I N S L L O Y F O N N R T L L L L T A 505
D B 445 D L Q I A I E N - - I T A S Y D N E T V T L Q D I C I A P S P Y N T - - - N C T I L S V L N Y F Q N S H S V L D H K K 499
QY 506 N Q T L M G Q T S V D W K D H E L Y C A N A P L T F K D G T A L A L S C W A D Y G A P V F F L A I G G Y K G K D Y S 565
D B 500 G D D F F - - - V T A D Y H T H E L Y C V R A P A S L N D I S L L H D P C L G T F G G P V F P W L V L G G Y D D Q N Y N 556
QY 566 E A E A L I M T S L N N Y P A G D P R L A Q A K L W E A P L E M R A P O R R M A G N F O V T F T A E R S L E D E I 625
D B 557 N A T A L V I T F P N N Y N D E K L Q R A Q A W E K E F I N V K Y N K N - - - E N L T I S T P A E R S I D E L 613
QY 626 N R T T A E D L P I P A T S Y I V I F L Y I S L A L G S Y S W S R V M V D S K A T L G L G V V A V L G A V M A A G 685
D B 614 N R E S D S D V F T V I S Y A I M F L Y I S L A L G H I K S C R L L V D S K V S L G I A G L I V L S S V A C S L G 673
QY 686 F F S Y L G I R S S L V I L Q W V P F L V S G A D N I F I V L E Y Q R L P R R P G E P R E V H I G R A L G R V A P 745
D B 674 V F S Y I G L P L T I V I E V I P F L V A G V D N I F L V Q A Y Q R D E R L Q G E T L D Q Q L G R V L G E V A P 733
QY 746 S M L L C S L S E A T C F F L G A L T P M P A V R T F A L T S G L A V I L D F L L O M A F V A L L S D S K R O A S 805
D B 734 S M F L S S P E T V A F F L G A L S V M P A V H T F S L P A G L A V F I D F L L Q I T C F V S L L G L D I K R Q B K N 793
QY 806 R L D V C C C K P Q E L P P P Q - G E L L L G F F Q K A V A P E L L H W I T R G V V L L F A L F G V S L Y S M 864
D B 794 R L D I P C C V R G A E D G T S V A S E S C L F R F F K N S T S P L L L K D W R P I V I A L F V G V L S F S I A V L 853
QY 865 C H I S V G L Q E L A L P K D S Y L L D Y F L P L N R Y F E G A P V Y F V T T L G Y N F S E A G N A I C S S A G 924
D B 854 N K Y D I G L Q S L S M P D D S T W D Y F K S I S Q Y L H A G P P V F V L E B G H D Y T S K G Q N M V C G G M G 913
QY 925 C N N F S F T Q K I Q V A T E F F P Q S Y L A I P A S S W V D D F I D W L T P - S S C C R L Y I S G P N K D K F C P S T 983
D B 914 C N N D S L V Q Q I F N A A Q L D N Y T R I G F A P S S W I D D Y F D W K P Q S S C R V - - - D N I T D Q F C N A S 970
QY 984 V N S N C L K N C M S I T - M G S V R P S V E Q F H Y L P W F L N D R P N I K C P K G L A A Y S T S V N - - L T S 1040
D B 971 V V D P A C V R - C R P L T P E G K Q P Q G D F M R F L P M F L S D N P N P K C G K G H A A Y S S A V N I L L G H 1029
QY 1041 D G Q V L A S R F M A Y H K P L K N S Q D Y T E A L R A A R E A L A N I T A D L R K V P G T D P A F E V F P Y T I T N V 1100
D B 1030 G T R V G A T Y F M T Y H T V L Q T S A D F I D A L K K A R L I A S N V T - E T W G I N G S - - A Y R V P P Y S V F Y V 1086
QY 1101 F Y E Q Y L T I P E G L F M L S C L V P T F A V S C L L G L D L R S G L L N L L S I V M L V D T V G F M A L W D 1160
D B 1087 F Y E Q Y L T I I D D T I F N L G S L G A I F L V T W V L L G C E L W S A V I M C A T I A M V L N M F G V M W L M G 1146
QY 1161 I S Y N A V S L I N L V S A V G M S V E F V S H I T R S F A I S T K P T W L E R A K E A T I S M G S A V E A G V A M T N 1220
D B 1147 I S L N A V S L V N L V M S C G I S V E F C S H I T R A F T V S M K G S R V E A E A L A H M G S S V F S G I T L T K 1206
QY 1221 L P G I L V L G A K A Q L I Q I F F R L N L L I T L L G L H G L V F L P V I L S Y V G P D V N P A 1272
D B 1207 F G G I V V L A F A K S Q I F Q I F F R M Y L A W L L G A T H G L I F L P V L L S Y I G P S V N K A 1258
RESULT 24
US-10-756-731-4
; Sequence 4, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstee, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894

```

; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1319
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-208-731-4

Query Match      34.5%  Score 2385; DB 4; Length 1319;
Best Local Similarity 40.0%; Pred. No. 7.9e-198;
Matches 524; Conservative 231; Mismatches 440; Indels 116; Gaps 28;

QY 14 LLLRLAQSEPYTHHQBGCYAFYDECGKNPELGSMLTSLNSVCLSNTPARKITGDHLIL 73
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 53 LLLLLCPAQFESQ-----SCWYGECC---IATGD---KRYNCKYSGPPKPLPKDGYL 100

QY 74 LQKICRPLYTGPNTOACCSAKOLVSLSEASITKALLTRCPACSDPNFVNHLCHNTCSNQ 133
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 101 VQSLCPGLFF--DNVSLCCDDIQQQLTKLSNLQLQLFLSRCPFCFYNLMTLFCBLTCSPHQ 159

QY 134 SLFINTRVAQLGAGOLPA---VVAEAPYQHSFAEQSYDSCSRVRVPAATAVLGWTMC 189
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 160 SQFLNATATEDYDPKTPENKTNVKELEYVQSGFANAMTNACRDVEAFSSNEKALGLLC 219

QY 190 GYVGSALCNAQRWLNFGQDTGNGLAPLDI-----TFHLLFPGAQVSGIOPLNEGVARCN 244
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 220 GRDARA--CNATNWEYFNKNDGQAPPTIIPVFSLSIL-----GMEPMENATKGN 270

QY 245 ESQGDVATCSCDCACPAIARPAQALDSTFYLGOMPSLVLIILICSVFVAVTILLVG 304
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 271 ESDEVETGPGSCQDCQIVCG--PKPQILQIGGGGW--GLDAMYVMWVTVYAFLEVFVG 327

QY 305 FRVA-----PARDKSKMVDPKGTSLSDKLSFSFTHLLGQF 340
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 328 ALLAVNCHRRRYFVSEYTPIDSNIAFSVNSDDGE-----ASCCDPLGAAPDDCLRM 380

QY 341 FQCGTWVASWPLTILVSLVPIVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHDOH 400
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 381 FTKWGAFCVRNPTCIIFSLAFITVCSGLVFQVTTNPVELMSAPNSQARSEKAFHDOH 440

QY 401 FGPFTNQVILTAPNRSRYRDSLLGPK--NPSGILDLDLLELLELQERLHLQVWSP 459
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 441 FGPFRTEQLIIQAPNTSVHIEPYEPAGADVPPGPPPLNKEILHOVLNLQ-----IAIESI 495

QY 460 EAQEN---ISLQDICVAPLNPDNTSLYDCINSLLQYFQNNRTLLLTANTQTLMGQTSQV 516
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 496 TASYNNETVLDICVAPLSYPNK---NCTIMSVLNYFQNSHAVL-----DSQV 541

QY 517 -----DWKHFLYCANAPLTFKDGATLALSCMADYGAPVPPFLAIGYKGYKDYSEAE 568
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 542 GDDFYIADYHHTFLYCVAPASLNDTSLHGPCLGTFGFPFVPLVLGYDDQNNAT 601

QY 569 ALIMPTSLNYPAGDPLRLQAOKLEWAEFLAEEMRAFORRMAGMFQVFTTAERSLDEINT 628
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 602 ALVITPPVNNYNDTERLQRAWAKEBFISFVKVKNYK---PNLTISFTAESRIDEELNE 658

QY 629 TAEDLPFATSYIVIFLYSLALGSYSSSRVWVDSKATLGLGGVAVTVLGAAMAAGFFS 688
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 659 SNSDVFVTIISYVVMFLYISLALGHIOQSCRLVDSKISLGIAGLIVLSSVACSIGIFS 718

QY 689 YLGIRSLVILQVVPVLSVAGADNIFVLEVYQRLPRPGRPREVHIGRALGRVAPSL 748
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 719 YMGMPITLVIEVIFPLVAVGVNDNIFILVQYQYQYQYQYQYQYQYQYQYQYQYQYQY 778

QY 749 LCSLSEACFFGALTTPMPAVRFTALTSGLAVILDFLLQMSAFVALLSLDSKRQASRLD 808
```

```

Db 779 LSSFSETSAFFPGALSSMPAVHTFSLFAGMAVLIDFLLOITCFVSLILGLDKROEKHLD 838
QY 809 VCCVVKPQELPPPOG---EGLLGFOKAYAPFLLHMITRQVWVLLLELFLALFVGLSYLM 864
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 839 ILCCVRGAD---DQGSASHASEYLFPPFKNYFAPLLKOWLRPIVAVFVGVLSFSAVV 895
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYEVEGAPVYFVTTLGYNFSSEAGMAIACSAG 924
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 896 NKVDIGLDQSLSPNDSYVIANFKSLAQYLSHSGPPVYFVLEEGVNYSSRKQGNVCGMG 955
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 925 CRRPSTQKIQYATEPEQSYLAIPASSWDDIDMLTP--SSCRLYISGNKDKPCPST 983
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 956 CNDLSLVOQIFNAAELEDTYTRVGFAPSSWIDYDFWVSPQSSCRLY---NVTHQFCVAS 1012
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 984 VNSLNLCKNCMSIT--MGSVRPSPVEQPHKLPWFLNDRPNIKCPKGGGLAAAYSTSVLTSDG 1042
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1013 VMDPTCVR--CRPUPGKQRPQKQBFMKFLPMFLSDNPNPKCGKGGHAAVGSANVIGDD 1071
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1043 QVL--ASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAPEVFPVYITNVF 1101
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1072 TYIGATVFWYTHITLKTADYTDAMKARLIASNITETWRS--KGSD--YRVFPYSVFPV 1128
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1102 YEOYLTPGLPMLSLCLVPTFAVSCLLILGLDLRLSGLNLNLSIVMLVDTVGFPMALWDI 1161
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1129 YEOYLTIIDDTIFNLVSLSGIFLVLTVLVLGCELWSAVIMCITIAMILVNMFGVWLWGI 1188
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1162 SYNANSLINLAVSGMSVEFVSHITSPALSTKPTWLERAKEATISNGSAVPAVAGVANTL 1221
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1189 SLNAVSLVNLVMSCGISVEFCSHITRAFTMSTKGSVRABEALAHMGSSVFSGITLTKF 1248
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
QY 1222 PGLVLGLAKAQIQLFFFLNLLIITLGLLHGLVFLPVLISVYVGDVNEA 1272
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1249 GGIVLAFAKSQIPFIFYFWYLAWLWGLGATHGLIFLPLVLLSYIGFSVAKA 1299
   ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :

```

RESULT 25

```

US-11-097-143-12003
; Sequence 12003, Application US/11097143
; Publication No. US2005020858A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-11-097-143-12003
```

Query Match	32.7%;	Score	2262;	DB	6;	Length	1287;
Best Local Similarity	37.8%;	Pred. No.	4e-187;				
Matches	504;	Conservative	244;	Mismatches	469;	Indels	116;
Gaps	30;						
Qy	10	LLWALLRLAQSEPTTHIQPCYCAFYDECGNPELSGLMTLSNVSCLSNTPARKITGD	69				
Db	17	LTAALVFLTIQSSKQD-----CWYGVGNTN-DFSHS-----QNCYPNGTAKEMATD	62				
Qy	70	HLILQKICPRLYTGTQACCSAKQLVSLASLSITKALLTRCPACSDNDFVNLHCHNTC	129				
Db	63	GLELLKRCGFLLENSENKFCDDKNQVELLNKNVELAGNILDRCFSCMENLVRHICQPTC	122				
Qy	130	SPNQSLFINVTRVAQIGAGQLPAVVAYEAFYQHSFABOSYSCSRVRPAAATLAVGTMC	189				
Db	123	SPKQAEFMHVATQNKXGD-EYISSVDLHISTEINKTYKSCQSVQPTGQLAFLDMC	181				
Qy	190	GUYGALCAQWALFQDGTGNGLAPLDTTHLLEPGQAVGSGIOPLNEGVARCNEQCD	249				
Db	182	GAYSASRCNPTKWFNFMGDATNYPYFQITIQHEP-KSNSNFTPLNVTTPCQNAVSS	240				
Qy	250	DVATSCODCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSVPFV-VT	299				
Db	241	KLPACSCSDCLSCQGPPEPRPPEPKIVGLDAYF-----VIMAAVFLVGL	288				
Qy	300	ILLVG---PRVAPARDKSNVDPKKGTSLSDKLSFS-----THLLGQFFQW	344				
Db	289	VFLMGSFLFTQSSMDNFQVD---GNDVSDEMPXSENDSYPEKLGHAHTETLETFFTK	345				
Qy	345	GTWVASMPLTILVSVIPVVALAAGLVFTELTPVLEWSAPNSOARSKAPHDQHPGF	404				
Db	346	GTFFASNPGLTITAGASLVILGYGINFTEITDPVKLWASNSKRLEREFDTKFSFP	405				
Qy	405	FRNQVILTAPNRSSVRYDSLLGPKNFSGIILDLLLELLELQERLRLQWSPSAQRN	464				
Db	406	YRLEQIIIAKAVLPOIVHNT-SNGPYTFGPVDFREFLTKVLDLQEGIKEINA-----NG	458				
Qy	465	ISLQDICYAPLNDNTSL--YDCINSLLOYFQNNRTLLLLFANQTLMOQTQVDWKDHP	522				
Db	459	TQLKDICYAPLSDGSEIDVSCVQSIWGYGDDRL-----DDHDEDNFNVYLDAL	514				
Qy	523	LYCANAPLTFKQGTALALSCMADYGAPVFPFLAIGVY-----KGKDYSEAEALIMTF	574				
Db	515	YDCISNPLY-----CLAPYGGVDPDAIALGGLFPDQDQLTGSTKFLANALITF	564				
Qy	575	SLNYPAGDPRLAQAKLWEAFLEEMRAF-QRMAGMFQVTTFAERSLEDEINRTAEDL	633				
Db	565	LVKQH-HNKTDLENALTWEKKFVEFTWNTYTKNNSQYMDIAFTSERSIEDELNRESQSDV	623				
Qy	634	PIFATSYIVIFLYISIALGYSYSSWSRVMDSKATLGLGVAVVLGVAMAMGFFSYLGIR	693				
Db	624	LTILVSYLIMFMYIAISLGHKFKRVFIDSKITLGTIGGVIIIVLASVSVSGVFGYIGLP	693				
Qy	694	SSLVLOVPFLVLSVGADNIFIFVLEYORLPRRGPPEPRVHIGRALGRVAPSMLLCSLS	753				
Db	684	ATLIIVEIFPLVAVDNIIFILVQHQDQKRPNETLEQQVQRIILGKVPMSLITLSL	743				
Qy	754	EATCFEFGALTPMPAVTFTALTGLAVILDFLLQMSAFVALLSDSKRQEARLDVCCCV	813				
Db	744	ESPCFFLGGISDMPAVRAFALYAGVALIIDFLQITCFVSLFTLTDKREENRMDICCFI	803				
Qy	814	KPOELPPPGQGGELLLGFFQKAYAPFLHLHWTITGVVLLLFALPGVSLYSNMCHISVLQDQ	873				
Db	804	KGKKPDSITNEGLLYKFFSVVYVFLMKKIVRASVMVIFPAMLCFSAIAIPRIDGLQD	863				
Qy	874	ELALPKDSYLLDYFLNRPFEVCAPVYFTVTLGYNFSSSEAGWNAICSSAGCKNFSPTQK	933				
Db	864	ELAMPQDSFVLHYFQSLNENLNIQPPYFVLKGLDAYTNSDQNLVLCAGQVCHNDSDVLQ	923				
Qy	934	IQVATEPPEQSYLAIAPASSVDDFDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNCL	990				
Db	924	IYLASRHSNQYIARPASSWIDDFDWAANAASSCKY-----RKDSGDFCPH--QDTSCL	976				
Qy	991	KNCMSITMGS--VRPSVEQFHKYLPWFLNDRNRIKCPKGGIAAAYSTSVNLTSDDQVL---A	1046				

Db	977	R--CNITKNSLLRPEKEFVKYLPFLKNDPDDTCARAGHAAYGAVRYNSHERLNTEA	1034				
Qy	1047	SRFMAHKPKNSQDYTEALPAARELANITADLR-----KVP-GTDPAPFEVPPYITN	1099				
Db	1035	SYPMAHYTLKSSADYFLALESARKISANITQMLQGRLMNSGVPMASALTVEVFPYSVY	1094				
Qy	1100	VFEYOYLTPLEGFMLSCLIVPTFAVSCILLGLDRLSLNLLSIVMLVDTVUGFMALW	1159				
Db	1095	VFEYOYLTMMSDITQSMGISVLSIFVTFVLMGDFVHSALVVVITITMIVNLGGLMYW	1154				
Qy	1160	DISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVMT	1219				
Db	1155	NISLNVSLNVLNVAWGLSVFECFCHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT	1214				
Qy	1220	NLPGTILVGLAKAQIQLIFFPRMLNLLITLGLLHGLVFLPVILSVYGVDPVNPALAEQKR	1279				
Db	1215	KFAGILVLAFAKSIQFQVYFRMYLGIIVVIGAAHGLIFLPVLLSYIGAPVSNARLRYHSQ	1274				
Qy	1280	A----EEAARAAM	1288				
Db	1275	AAAHETALAGIL	1287				
RESULT 26							
US-11-097-143-2679							
; Sequence 2679, Application US/11097143							
; Publication No. US20050208558A1							
; GENERAL INFORMATION:							
; APPLICANT: Venter, J. Craig							
; APPLICANT: et al.							
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID							
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE							
; TITLE OF INVENTION: DROSOPHILA GENES.							
; FILE REFERENCE: CL000728							
; CURRENT APPLICATION NUMBER: US/11/097,143							
; CURRENT FILING DATE: 2005-04-04							
; PRIOR APPLICATION NUMBER: 60/157,832							
; PRIOR FILING DATE: 1999-10-05							
; PRIOR APPLICATION NUMBER: 60/160,191							
; PRIOR FILING DATE: 1999-10-19							
; PRIOR APPLICATION NUMBER: 60/161,932							
; PRIOR FILING DATE: 1999-10-28							
; PRIOR APPLICATION NUMBER: 60/164,769							
; PRIOR FILING DATE: 1999-11-12							
; PRIOR APPLICATION NUMBER: 60/184,831							
; PRIOR FILING DATE: 2000-02-24							
; PRIOR APPLICATION NUMBER: 60/191,637							
; PRIOR FILING DATE: 2000-03-23							
; NUMBER OF SEQ ID NOS: 43008							
; SOFTWARE: FastSeq for Windows Version 4.0							
; SEQ ID NO 2679							
; LENGTH: 1223							
; TYPE: PRT							
; ORGANISM: DROSOPHILA							
US-11-097-143-2679							
Query Match							
Best Local Similarity							
Matches							
27.0%; Score 1864.5; DB 6; Length 1223;							
33.5%; Pred. No. 1.7e-152;							
Conservative 237; Mismatches 498; Indels 113; Gaps 27;							
Qy	62	PARKITG-DHILLQKICPRLY-----TQPN-TQACCSAKQLVSLASISITKALLTRC	113				
Db	14	PAPLNSPTSEATFAKRCPMLYKEYKGESGEDELSCDAAQIETWESGLSQADGVFSRC	73				
Qy	114	PACSDNFVNLHCHNTCSFNOSLFINTVRAQLGAGQLPAVVAYEAFQHSFAEQS----	169				
Db	74	PTCTRNALTVCACTCAKQHTLFLTAYNDTN-----DAGVDYVKYIDYRLTDDTVSKIY	127				

170 DSCSRVRPAAATLAVCTMGVYGSALCNQORWLNFGDGTNGGLAPLDTITFHLLEPGQAV 229
128 NSCIGIQHTQGTGRPMDLGGCSYNAKTCNRRRYTFPMGDSYGVNPPQINRYKWSDEABE- 186
230 GSGIQPLNEGVARNCSQGDVATCSQDCAAPAIARPAQALDSTFYLGQMPG--SLVL 287
187 GSNEIYLDLSPKGGESYDSA-CACIDEECSPLTDAFTGDBELWKIAGLIGVTFILA 245
288 IILCSVFAVVTTLLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQF----FOG 343
246 LILACALSFFIFWGAFTKSAP-----SVCMTPLFGBFFYHGFR 285
344 WGTWVASWPLTILVLSVIPVVALAAGLVFTELTTDPVELWASPNQARSEKAPHDQHPG 403
286 WGFPCARHPVIVLALCSWLAGSFGIRYNTITDPVELWAGESOTRIKQDFDQHPG 345
404 PFTNQVILTAPNRSSRYDLSLLGPKNFSGIIDLDDLLLELLELQERLRLHLOVMSPEAQ 463
346 FYRTNQMFVAVNQTYFTHET-SNGVNLFGPAFEYNPLKEVPELQDSIMKLG-AD 399
464 NISLQDICVAP-LNPONTSLYD--CCINSLLQYFQNNRTLLLLPANQTLMGQTSQVDMKDH 521
400 NEGLDKICVAPVLMAGETPTVDRCAIQSVVGYFQHDMDRF--ENSVDNNNTYINLQ 456
522 FLYCANAPLTFKDTALASCMADYGAPVPPFLAIGY-----KGDYSBAAALWTFPS 575
457 LEDCLRVP-MED-----CFPTGCGPIEPGIAVGMPKVAVGEDPDYMLATGLVLTF 508
576 LNNYPAGDPLAQAKLWBAFLBEMAFQRMAGMFQVTFABERSLEDEINRTAEDLPI 635
509 GRNY-NDESKLEPMKWEKLVDFLDRYK---SDRLDIATMAERSIQDAIVELSEGEVST 564
636 FATSIVIVILYISLALGYSXSSWRVMVDKATILGAGVAVVLGAVMAAMGFFSYLGIRSS 695
565 WLSYVVMFYVAIALGHIRSCRGFLRESRIMLAIGGIIVLASVVSIGLFGWGLDVTYT 624
696 LVTLQVPELVLSVGNADNIPFVLEYQRLPRRPGEPREVIHGAIRGAPVAPSMMLCSLSA 755
625 MIAIEVPELVLAAGVDNIFIMVHTYQRLDHSKFKTTHEAIGBAIGVQGPSILQTAGSEM 684
756 ICFPLGALTPMPAVRTFALTSLGLAVLDFLQMSAFVALLSDLSKQESRLDVCCKV- 814
685 ACFAICISDMPAVKTFAMATAIILDFLQITAFVAMALDEKRYLDRGLDMLCCKVS 744
815 -----PQELPPPCQGGELLLGFQKAYAPFLHWTIRGVVLLILFLALFGV 859
745 GKGKINDEGDGVDREKVE-----GLLETFLXNFYSPFLSKPKVSVLLIFTITCL 797
860 SLYSMCHIISVGLDQELALPKDSVLLDYFLFLNRYFVGAPVYFVTTLYGNFSEAGMNAI 919
798 SLWVTPSIEKGLDQEMSPKNSHVVKYFRYVMDLLANGAPVYVWLKPLNYSPLQONLI 857
920 CSSAGCNNSFPQKIQYATEFPEQSLAIPASSWDDPDWLTSPSCCHLYISGPNKDXP 979
858 CGGVECNNSLSVQLYTOAQYBEITSARPASSWLDYIDWLAISDCCRYNT---TGGF 914
980 CPSTVNSLNLCKNCSMTGMSVRPSVEQFKYLPWFLNDRPNIKCPKGLAAYSTSVNLT 1039
915 CSNSKSEBCLPCRGETENGRLPDAETFNKIPIYFLPLDPAEACKAGASTADAVIYT 974
1040 SD----GOVLASRFMAYHKLKNSQDYTBALRAARELAANITADLRKVPGTDPAPFVFPY 1095
975 IDDVGMSTVQDSYFMQYSTSTTSBEEFYSQLRVREIRISGEINAMPKE---NNVDABIFAY 1031
1096 TITNVEYQYILPGLFMLSCLVPTFAVSCLLGLDLRSLGLMLLSIVMLVDVTGQF 1155
1032 CVFYIYEQYILWGDAMFSLGMSLVAIEFLVTLITGLDITSTFVFLMVICILINMLGM 1091
1156 MALWDSYNAVSLINLVSAGVMSFVSHITSEFALSTKPTWLBRAKEATISMGSAVPAG 1215
1092 MWANSINLNAISLVNVCVGVGEFVAHIVRSFK-RAEGTAQEARHSLNTGSSVLSG 1150
1216 VAMTNLPGILVLGLAKAQIQQIFFFLNLILITLGLLHGLVFLPVILSVGPDVNPALAL 1275

1151 ITLTKFAGIVVLGFSNSQIFQVFFRMYLIGVILGAAGHLILLPVLLSLGPPQK----L 1206
1276 EOKRAEEAAVAVMVAS 1291
1207 ARSSGABPTASITITT 1222

RESULT 27

US-10-208-731-6
; Sequence 6, Application US/10208731
; Publication No. US20030092038A1
; GENERAL INFORMATION:
; APPLICANT: Carstea, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; CURRENT FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1170
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-208-731-6

Query Match 19.2%; Score 1329; DB 4; Length 1170;

Best Local Similarity 28.8%; Pred. No. 8.1e-106;

Matches 373; Conservative 230; Mismatches 512; Indels 182; Gaps 46;

10 LLWAL-----LLRLAQSEPTTTHQPGYCAFYDECGKNP-----ELSGSLMTLSNVSCLSN 60
3 LLIILAVGQALMLVCG-----TATCAMYGNCCKSVFGNELPCVPSPSEPPVLS 54
61 TPARKITGDHLILLOKICPRLVTGNTQACCAKQVLSBASLSITKALLTRCPACSDNF 120
55 ETSK-----LLVEVCGEWEKVR-YACCTKQVVALRDNLQKAQPLISSCPACLKNF 105
121 VNLCHNTCSNOSLPINVTVAQLGAGQLPAVVAEAFVQHSFASQSDSCSRVVPAA 180
106 NMLFCHTCAQGGFVNITKV-EKSKEDKIDVAELDVFNWSSWASEFTDSCNIIKPSAT 164
181 ATLAVGTMCGVYGSALCNQORWLNFGDGTGN--GLAPLDTITFHLLEPGQAVSGIQPLNE 238
165 N----GYAMDLLGGGAKNYSQFLKFLGDAKPMLGSGSPQINRYKYDLANEE--KEWQEFND 218
239 GVARNESGDDVATCSQDCAAPAIARPAQALDSTFYLGQMPG---SLVLIILCSVP 295
219 EYVACDDAQ----YKACSDCQSCPHL-KP-LKDGVCVKVGPLCFSLSVLIFYTTICALF 272
296 AVVTILLVGFVRVAPARDKSMVDPKKGTSLSDK-----LSFSTHT-----LLCQPFQCGW 345
273 AFMYTYLCKRKNKGAMIVDDDDIVPSSG-SLDESETNVFESFNNETNPFNGKLANLFTKVG 331
346 TWVASWPLTILVLSVIPVVALAAGLV-FTELTTDPVELWASPNQARSEKAPHDQHPGF 404
332 QFSVENPYKILLITTVFSIFVFSFIIPQYATLETDPINLWVSKNSEKFEKEVFDNFGFP 391
405 FRTNQVILTAPNRSSRYDLSL--LLGPKNFSGIIDLDDLLLELLELQERLRLHLOVMSPEAQ 462
392 YRTEQIFVFNNETGTVLSYETLHWFDVENF-----ITEEL-----QSS 429
463 RNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLLPANQTLMGQTSQVDMKDH 522
430 ENIGYQDLFRP-TEDST----CVTESPTQYFQG-----ALPNKDSWKREL 470
523 LYCANAPLTFKDTALASCMADYGAPVPPFLAIGYKGYKGYSEABALWTFPSLNPAL 582

Db 879 VYVHDPN--TFCSTNRKNSALDDKACKTCMDP----- 908
QY 1025 PKGLAAYSTSVNLTSQGVLASRFMAYHKPL--KNSQDYTEALRAARELAANITADLRK 1082
Db 909 --DGRASFDAISFTRGRQASQFMTFHKKLSISNSSDFIKAMDYARWYSRRLERSI-- 964
QY 1083 VPGTDPAFVPPYTIITNVFYEQYLTILPEGLFMLSCLVPTPAVSCLLGLDRLSLNL 1142
Db 965 ----DDTAHFVAYSFKIPFPFYEQYSTIMPILTQLFITVGVFGIICVTGLIDVKGACAV 1020
QY 1143 LSIIVMLIVDTVGPMAWDISYNVLSNLVSAGVMSVEFVSHITRSFAISTKPTMLERAK 1202
Db 1021 ICQVS-----NYFHVSS--GILIEFVNVLKGFACSLRQAKDRAE 1060
QY 1203 EATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIPFFRLNLLITLLGLHGLVFLPVIL 1262
Db 1061 STVSGIPIILSGPVVMTAGSTWFLSGAHLQIITVYFFKLFLITIVSSAVHALIILPILL 1120
QY 1263 SYVGPDVNPALALEQKRAEAAVAVMASCNPHSPRSVSTADNIYV-----HSFEGS 1314
Db 1121 AFGSGRHGSGSETSTNDNDQHDACVLS--PTAESHSINVEEGILNRPSSILLDASHILDPL 1178
QY 1315 IKGAGAISNPLNNGROF 1332
Db 1179 LKAEGGIDKAI-----GRDF 1193

RESULT 29

US-10-239-316-8
; Sequence 8, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: KITAYAMA, Yoshio
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USP
; CURRENT APPLICATION NUMBER: US/10/239, 316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 8
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-8

Query Match 13.6%; Score 942; DB 4; Length 194;
Best Local Similarity 99.5%; Pred. No. 2.7e-73;
Matches 193; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 615 FTAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSYSWSRVMDSKATLGGVA 674
Db 1 FNAERSLEDEINRTAEDLPFATSYIVIFLYISLALGSYSWSRVMDSKATLGGVA 60
QY 675 VVLGAVMAAGFPESYIGIRSSVLQVLPVLVLSUCADNIFIVLEYQRLPRPGPREV 734
Db 61 VVLGAVMAAGFPESYIGIRSSVLQVLPVLVLSUCADNIFIVLEYQRLPRPGPREV 120
QY 735 HIGALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 794
Db 121 HIGALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 180
QY 795 LSLDSKQKQASRLD 808
Db 181 LSLDSKQKQASRLD 194

RESULT 30

US-10-424-599-211862

; Sequence 211862, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211862
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33338C.1.pap
US-10-424-599-211862

Query Match 13.0%; Score 900.5; DB 4; Length 492;
Best Local Similarity 39.3%; Pred. No. 4.9e-69;
Matches 199; Conservative 99; Mismatches 170; Indels 39; Gaps 11;
QY 736 IGRALGRVAPSMMLCSLSAICFFLGCALTMPAVRTFALTSGLAVIDFLQMSAFVAL 795
Db 1 ISIALVEVGPSITLASVSEVLAPAVGFSFISMPAIRVFSMFAALAVLDFLQVTFVALI 60
QY 796 SLDSKQKQASRLDVCVCPQBLPPPGQE--GILLGPFQKAYAPFLHWTIRGVVLLL 852
Db 61 VLDSLRADKRVDCPCIKVHADPDIGTGRKPGLLARYMKVHAPILSIGVKIVIAI 120
QY 853 FLALFGVSLYSMCHISVGLDQELALPKDSYLDYFLFLARYEVGAPVYFVTTLGYNESS 912
Db 121 FVGFALASALTSTRIEPLGLEQIVLPDRDSYLOGYFNNVSEYLRIGPPVYFVVK-NTNYSS 179
QY 913 EA-GMNAICSSAGCNFSTQIQVATEPPEQSYLAIPASSWDDDFIDMLTPSS--CCRL 969
Db 180 ESTHTNQLCSISHCNSDSLNEIARAALVPDTSYIAKPAASWLDLFLVWSPFAFGCCRK 239
QY 970 YISGPNKDKFCP-----STVNSLNCNKMCSI---TMGSVRPSVEQFHKYP 1013
Db 240 FTNG----SYCPDDQPPCCAPGSSCVSGTKDCTTCFRHSDLHNDPSTTQPREKLP 295
QY 1014 WFLNDRPNIKCPKGLAAYSTSVNLT--SDGQVLASRFMAYHKPLKNSODYTEALRAARE 1071
Db 296 WFLSLSPSADCAKGGHGAYTSSVELKGYDNGIIOASSFRYHTPLNKNQDYVYNSMRAARE 355
QY 1072 LAANITADLRKVPGTDPAPEVFPYITNVFYEQYLTILPEGLFMLSCLVPTPAVSCLLL 1131
Db 356 FSRVSDSLK-----IEIPYSVYFMYFPFQYLIHWKTALINLAIGAIVIV-CLIP 406
QY 1132 GLDLSGLNLISIVMLVDTVGFPMALMDISYNAVSLINLVSAGVMSVEFVSHITRSFAI 1191
Db 407 TSSLSWSSSIILLVAMIVVDLMGLMAILNIQLNALSVNLVMSVGIAGEVFCVHMTSHFTV 466
QY 1192 STKPTWLERAKETISMGSAVFAGVAM 1218
Db 467 ASGDR--DORAKEALGTMGASVFSGITL 492

RESULT 31

US-10-424-599-189288
; Sequence 189288, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 211862
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_33338C.1.pap
US-10-424-599-211862

```
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 189288
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_141944C.1.pap
US-10-424-599-189288

Query Match          12.4%; Score 859.5; DB 4; Length 541;
Best Local Similarity 35.3%; Pred. No. 2.1e-65;
Matches 197; Conservative 109; Mismatches 187; Indels 65; Gaps 14;

QY 752 LGEACFFGLGTPPAVTEALTSLGLAVILDFLLQMSAFVALLSLD--SKRQESRLDV 809
DB 2 LSCRVC-----SMIPAPAFRIY-----YIPLITAFVALITFDIFISAYMESFAPFP 46
QY 810 CCCVKPQELPPQCGE-----GLLGFFQKAYAPFLHWTIRGVVLLFLALFGVS 860
DB 47 CVMLEP---PSAERNEGTCVRRERDGLITRYMKVEHAPFLGLCGVKILVIAVFAAFTLAS 103
QY 861 LYSMCHISVGLQELALPKDSVLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSEA-GMNAI 919
DB 104 IALCTRIEPLGEOQIALPDSYLGQYFSNISYLRVGPPLYFVVK-DYNISLESKHTNQL 162
QY 920 CSSAGCNNSFTQIKYATFEPEQSYLAIPASSWDDFDLWLTPTS--SCCRLYISGPNKD 977
DB 163 CSISHCDSNLSNEISRAISLVPTSSYIAKPAASWLDLDFLWISPEAFSCCKXF---TND 218
QY 978 KFCPTSVNSLNCNKCMSITMGSV-----RPSVEQFHKYLFWFLNDRPN 1021
DB 219 SYCPDDQPPCCLPDPGCGGLGGCKDCTTCFRHSDLVNDRPSTAQFREKLEFWLDALPS 278
QY 1022 IKCPKGLLAYSTSVNLT--SDQVLASRFMAVHKPKNSQDYTEALRAARELAANITAD 1079
DB 279 ADCAKGGHAYTNSVDLNGYEGGVIOASEFRTYHTTVNQGDKVNAIRAROFSAKISS 338
QY 1080 LRKVPGTDAFAFVFPYITNFFVEQYLTILPEGLFMLSCLPFTFAVSCLLGLDLRSGL 1139
DB 339 LK-----WDIFPYSVFVIFEQYLDIWKLALINISALGAIFV-CLITSSVWSSV 389
QY 1140 LNLISVIMLVDTVGFMAWMDISYNAVSLINLVSAGMSVEFVSHITRSPALSTKKTWLE 1199
DB 390 ILLVLMIMILDLGMVMAILGIQLNAVSVVNLIMSIGIAVEFCVHIVHAFTVSLGDR-SQ 448
QY 1200 RAKENTISGSAVPAGVAMTNLPGILVGLAKAQLIQIPFRNLTLTLGLHLGLVFLP 1259
DB 449 RAKTALCTMGASVFSGITLTKVGVVLCTFSTSQIFVYIFQMYLALVILGLHLGLVFLP 508
QY 1260 VILSVYGPVDPNPALEQ 1277
DB 509 VVLSLFGPPLRYTVIKEQ 526

RESULT 32
US-08-954-701A-19
; Sequence 19, Application US/08954701A
; Publication No. US20030032085A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
```

```
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1447 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-954-701A-19

Query Match          9.4%; Score 651.5; DB 2; Length 1447;
Best Local Similarity 22.0%; Pred. No. 1.4e-46;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PQQAVSGIQPLNEGVARCNESQGDVATVCSQDCAASCPAIAIPAOLDSTFYLGQMPQS 284
DB 24 PQRPAAGRRRTTGLRRAAAPDRD-----YLHRPSYCDAAFALEQI--- 65
QY 285 LVLIILCILSVFNAVITILLVGRFVAPADKSKWVDPK-----KGTSLSDKLSFSHTLL 337
DB 66 LVLIILCILSVFNAVITILLVGRFVAPADKSKWVDPK-----SKGATGRKAPLWRAKFORLLFKLGYIQKNC 98
QY 338 GQFFQMGWGTWASWPLTILVLSVIPWALAAGLVFETLTDPELWSAPNSQARSEKAPH 397
DB 99 GRP-----LVVGLLIFGAPAVGLKKAANETNVEELVVEGVGVRVSRLENT 143
QY 398 DQHPGFFFRTN-QVILTAPNRSRYSDLSLLGPKNFSGLDLDLLELE--LQERLRHL 454
DB 144 RQKIGEEAMFNQMLQTPKEBG-----ANVLTEALLQHLDSALQASRVHV 190
QY 455 QWVSEAPQRNLSLDICYP-----LNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
DB 191 YMYN-----RQWKLEHLCKYKGBLITETGYMDQIIEYLYPCLITITLDCFWEGKLOSCTA 246
QY 506 NOTLMGQ-----TSQVDWKDHFY-----CAN----- 527
DB 247 ---YLGGKPLRWNTNPDLEFLBELKKINQVDSWEEMLNKAEVGHGMDRPCLNADPDC 304
QY 528 -----APL-----TFKQGT-----ALALSCMA 544
DB 305 PATAPNKNSTKPLDMALVNLGGCHLSRKYMHWQEBELIVGGTVKNSTGKLSVAHALQTMF 364
QY 545 DYGAVPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLQAQAQLWEAFLEENRAFQ 604
DB 365 QLMTPKQMY---EHFKGYEY-----VSHINNE-----DKAAALLEANQRTYVEVHVQSV 411
QY 605 RRMAGMFQVFTTAERSLEDEINRTTAEDLPFATSYIVIFLYTSLALGSYSSRSRVMVDS 664
DB 412 AQNSTQKVLSTFT--TTLDDILKSFSDSVIRVASGYLLMLAYACTLM---LRWD--CSKS 465
QY 665 KATLGLGGVAVLVGAMAAAGPFSYLGIRSSLVILQVPPFLVSVGNADNIFIFVLEVQRL 724
DB 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAAATQVLPFLALGVGVDDVFLLAHAFSET 525
QY 725 PRPGEPRVHIGRALGRVAPSMGLCSLSAICFFLCAITPMPAVRTFALTSLGLAVLIDF 784
DB 526 GQNKRIFFEDRTGECLKRTGASVALTISNVTAFMAALIPALRAFSLQAAVVVVVNF 585
```


Db	914	INPSA--FYIYLTAWNSNDPVAASQANTIRPHRPEWHDK---ADYNETRLRIIPAAEP	968
Qy	1044	VLASRFMAYHKPLKNSQDYTEALRAARELAANTITA-DLRKVPCTDPAFVFPYTIITNVFY	1102
Db	969	IEYAQFPFYINGLRDTSDFVEAIEKVRTICSNVTSLGLSSYPNG-----YFP---LFW	1018
Qy	1103	EQYLTILPEGLFMLSICLVTEFAVSCLLGLDLRSGLNLNLSITVMIIVDTVGPALWDLS	1162
Db	1019	EQYGLRHWLLFISVVLACTFLVC AFLNPTWAGII-VVVIALMIVELFGMGLIGIK	1077
Qy	1163	YNASVNLINLNASVGMSEYFESHITRSP--AISTKPTWLERAKEATISMGSAVFAGVAMTN	1220
Db	1078	LSAPVPVILLIASVGIGVEFTVHVAFALFLTAIGDKN---RRAVLALBHPFAPVLDG-AVST	1133
Qy	1221	LPGLTIVLGLAKAQLQIUFFEFLNLLITLGLLHGLVFLPVILSVYG--PDVNPALALEQ-	1277
Db	1134	LLGVLMIAIGSEDFIVRYFVAVLAILTILGVNLGVLLPVLLEFFGYPPEVSPANGLNRL	1193
Qy	1278	--KRAEEAFAAVMVASCPNHPSRVSTADNIYVNHSPFGSIGK	1317
Db	1194	PTPSPPEPPPSVVRFAMPFGHTH--SGSDSSDSSEYSSQTWSG	1233

RESULT 34

```

US-09-754-032-19
; Sequence 19, Application US/09754032
; Publication No. US20030148388A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; GOODRICH, LISA V
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001
; CLASSIFICATION: 435

```

PRIOR AE

```

1  APPLICATION NUMBER: US/08/540,406
2  FILING DATE: 06-OCT-1995
3  ATTORNEY/AGENT INFORMATION:
4  NAME: Rowland, Bertram I
5  REGISTRATION NUMBER: 20015
6  REFERENCE/DOCKET NUMBER: a60190-1
7  TELECOMMUNICATION INFORMATION:
8  TELEPHONE: 415-781-1989
9  TELEFAX: 415-398-3249
10 INFORMATION FOR SEQ ID NO: 19:
11 SEQUENCE CHARACTERISTICS:
12 LENGTH: 1447 amino acids
13 TYPE: amino acid
14 STRANDEDNESS: single
15 TOPOLOGY: linear
16 MOLECULE TYPE: protein
17 SEQUENCE DESCRIPTION: SEQ ID NO: 19:
18 US-09-754-032-19
19
20 Query Match 9.4%; Score 651.5; DB 3; Length 1447;
21 Best Local Similarity 22.0%; Pred. NO. 1.4e-46;
22 Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

```


Db 969 IEVAQPFYINGLRDTSDFEALIEKVKITCSNTYSLGLSSYPNG-----YFP-----LFW 1018
QY 1103 EQLYTLPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLISIVMILVDTVGFMALWDIS 1162
Db 1019 EQYIGLRHLLFISVVLACTFLVCAVFLNPNWTAGII-VNVLALMTVELFGMGLIGIK 1077
QY 1163 YNAVSLINLVSAGMSVEFVSHITRSP--AISTKPTWLEBAKATISMGSAGVAVMTN 1220
Db 1078 LSAVPVVILITASVIGVEFTVHVALAFLTAIGDKN---RRAVLALHMFAPVLDG-AVST 1133
QY 1221 LRGILVLGAKAQLIQIFFRRLNLLITLGLLHGLVPLPVILSVG--PDVNPALALEQ- 1277
Db 1134 LLGVMLAGSEFFIVRYFFAVIALITLGLVNLGLVLLPVLFFGYPPEVSPANGLNRL 1193
QY 1278 --KRAEEAANVMVASCPNHPSPVSTADNIIYNHSPSGSIKG 1317
Db 1194 PTPSPPPPSVVRFPMPGHTH--SGSDSDSEYSGTTVSG 1233

RESULT 35

US-10-421-446-19
; Sequence 19, Application US/10421446
; Publication No. US20030186309A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; GOODRICH, LISA V
; JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/421,446
FILING DATE: 22-Apr-2003
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/724,631
FILING DATE: 28-No. US20030186309A1-2000
APPLICATION NUMBER: US/08/656,055
FILING DATE: 1996-05-31
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Query Match

Best Local Similarity 9.4%; Score 651.5; DB 4; Length 1447;
Matches 300; Conservative 181; Mismatches 460; Indels 421; Gaps 47;

QY 225 PGAVSGIQPLNEGVARCNESQDDVATCSCQDCAASCPIARPOALDSTFYLGQMPGS 284
Db 24 PGRPAGGRRRTTGGRLRAAADPRD-----YLHRPSYCDAAFALEQI--- 65
QY 285 LVLIILILCSVFVAVVTILLVGRVAPARDKSKMVDPK-----KGTSLSKLSFSTHTLL 337
Db 66 -----SKGKATGRKAPLWLRKRAKFORLLFLKGLCYIQKNC 98
QY 338 GQPFQCGMTWASWPLTILVLSVIPVVALAAGLVFTTETDPVELMSAPNSQARSEKAFH 397
Db 99 GKF-----LVVGLLIIFGAFVAGLKAANLETNVEBELWVEVGRVSRRELYT 143
QY 398 DQHPGPFPRTN-QVILITAPNRSSRYDLSLLGPKNFSGLIDLDLLLELLE--LQERLHL 454
Db 144 RQKIGBEAMFNPQLMIQTPKEG-----ANVLTTEALLQHLDSALQASRVHV 190
QY 455 QWSPQAQRNLSLODICYAP-----LNPNSTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 191 YMYN-----RQWKLHLCKYKSGELITETGYMDQIIIEYLYPCLITPLDCFWEGAKLSGTA 246
QY 506 NOTLMGO-----TSQVDWKDHFY-----CAN----- 527
Db 247 --YLLGKPFIRWTNFPDLFLBELKKNINYQVDSWEEMLNKAEVGHGYMDRPCLNADPDC 304
QY 528 -----APL-----TPKCGT-----ALALSCMA 544
Db 305 PATAPNKNSTKPLDMALVNLGGCHGLSRKYMHWQBELIVGGTVKNSTGKLVSAAHALQTMP 364
QY 545 DYGAVPFPFLATGGYKDYSEBALIMTFSLNNYPAGDPRLAQAKLWEAPFLERAFQ 604
Db 365 QLMTPKQMY---EHFKGYEY-----VSHINWNE-----DKAAAILAEAWQTYVEVVHQSV 411
QY 605 REMAGMFOVFTABERSLEDEINRTAEDLPIPATSYIVIFLYISLALGSSVSSSRVWVDS 664
Db 412 AQNSTQKVLSTFT--ITLDDILKSFSDSVIRVASGYLLMLAVACLTM--LRWD--CSKS 465
QY 665 KATLGLGVAVVGLVGAAMAGFFSYLGRSSVILQVWPFLVLSVGADNIFIVLVSQRL 724
Db 466 QGAVGLAGVLLVALSVAAGLGLCSLIGISFNAATTOVLPFLALGVGDDVFLLAHAFSET 525
QY 725 PRPGEPREVHI GRALGRVAPSMMLCSLSEALCFIFLGALTMPMPAVRTALTSLAVILDF 784
Db 526 GQMKRIPFEDRTGECLRTGASVALTSISNVTAFMAALIPALRAFSLQAAVVVVVNF 585
QY 785 LLOMSAFVALLSLDSKQSEASRLDVCCC-----VKQEL-----PPPG 822
Db 586 AMVLLIFPAILSMDLRYRREDRLDIFCCFTSPCVSRVIQVBFQAYTDTHTNTRYSPPPY 645
QY 823 QGEGGL----- 827
Db 646 SSHSAHETQITMQSTVQLRTBYDPTHVYVYTABRSEISVQPVTVTQDTLSCQSPST 705
QY 828 -----LLGFFOKAYAPFLHMTITRGVLLLLFLALFGVSL 861
Db 706 SSTDRLLSQFSDSLHCLPEPCTKWTLSFAEKHYAPFLKPKAKVWVIFLFLGLIGVSL 765
QY 862 YSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA PVYFTVLGYNFSSESGMNAICS 921
Db 766 YGTTVRDGLDLDIVPRETREYDFIAAQKPYSP-----YNN----- 803
QY 922 SAGCNNSFTQKIQYAT-----EPPEQSYLAIAPASS-----WVDDFDIDLW----- 961
Db 804 -----YIVTQKADYFNIOHLLYDLHRSPSNVYVWLEENKQLPKWMLHYFRDWLQGLQD 857
QY 962 -----TPSS-----CCRLYISGPNKDKPCPSTVNSLNCNKCMSTWGS 1000
Db 858 AFDSDWETGKIMPNNVKNGSDGVLAYKLLVQTGSRDK--PIDISQLT--KQRLVDADGI 913
QY 1001 VRPSVEQPHKYL-PWFLND-----RPNKCPKGLAAY--STSVNLTSDDQ 1043
Db 914 INPSA--FYIYLTAWVNDPVAAYASQANIRPHRPEWHDK---ADYMPETRLRIPAASP 968

FILING DATE: 07-JUN-1996		954	VDDFIDWL	-----TPSS-----	CCRLYISGPNKDKFCSTVNSL	987
APPLICATION NUMBER: US 60/019,765						
FILING DATE: 14-JUN-1996		694	LHYFRDLQGLQDAFSDWETGKIWNPNYKNGSDGVLAYKLLVQTGRDK--PIDISQL	751		
ATTORNEY/AGENT INFORMATION:						
NAME: Hyman, Laurence J.		988	NCLKNCMSITMGSVRPSVEQFHKYL--PWFLND-----	RPNIKCPKGGGLAAY	1032	
REGISTRATION NUMBER: 35, 551						
REFERENCE/DOCKET NUMBER: 015280-278200US		752	T--KQRLVDADGIINPSA--FYIYLTAWVSNDPVAYAASQANIRPHRPEWVHDK--ADY	804		
TELECOMMUNICATION INFORMATION:						
TELEPHONE: (415) 576-0200		1033	--STSVNLTSQGVLASRFMAYHKPLKNSQDYETALRAARELANITA-DLRKVPGTDP	1089		
TELEFAX: (415) 576-0300		805	MPETRLRIPAAEPIEYAFPPFLNGLRDTDFVEAIEKVRTICSNYTSGLSSYPNG---	861		
INFORMATION FOR SEQ ID NO: 60:						
SEQUENCE CHARACTERISTICS:		1090	FEVFPYITINVFYBOYLITLPEGLFMLSCLVPTFAVSCLLGLDLGLSLMLLSIVMIL	1149		
LENGTH: 1296 amino acids						
TYPE: amino acid		862	---YFP-----LFWEQYIGLRHWLLFISVLACTFLVCANVLLNPWTAGII--VMVALMT	913		
STRANDEDNESS: <Unknown>						
TOPOLOGY: linear		1150	VDTVGFMAIWDISYNVSLINLVSAVMSVEFVSHITRSP--AISTKPTWLERAKEATIS	1207		
MOLECULE TYPE: protein						
FEATURE:		914	VELFGWGLIGIKLSAVPVVILIASVGIVETVHVVALAFLTAISDKN---RRVLALEH	970		
NAME/KEY: Protein						
LOCATION: 1..1296		1208	MGSVAFAGVAMTNLPGIILVGLAKAQLIQIFFPRNLNLLITLLGLLHGLVFLPVILSYVG-	1266		
OTHER INFORMATION: /notes "amino acids encoded by human						
nevold basal cell carcinoma syndrome		971	MFAPVLDG-AVSTLLGVMLAGSEDFIVRYFFAVLAILTLLGVNLGLVLLPVLUSFFGP	1029		
(NECCS) (PATCHED (PTC)) CDNA"						
SEQUENCE DESCRIPTION: SEQ ID NO: 60:		1267	-PDVNPALALEQ---KRAEEAFAVAVMVASCPNHPSPVSTADNIYVNHSEFGSIK	1317		
US-10-302-279-60						
Query Match						
Best Local Similarity						
Matches						
426		LLG--PKNPSGILDLLLELELELERHLQVWPEAORNIS-----LQDICVAPLNP	477			
97		LLGKPLRWNTNPFLEFLBEXKI-----NYQVDSWEELNKAEEVGHGYMDRPNCLNPADP	151			
478		DNTSLYDCCINSLLQFQNNRTLLLTANOTLMTGQTSQ--VDMKDHFLYCANAPLTFKQGT	536			
152		D-----CPATAPNKNSTKPLDMALVNGCGHGLSRKYHWHQBELIVGG-----TVKNS	200			
537		-----ALALSCMADYGAPVFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQA	591			
201		GKLVSAHAAQTMLTPKQMY---EHFKGYEY-----VSHINWE-----DKAAAL	247			
592		WEBAFLLEEMRAFORMAGMQVTTAERSLEDEINTTAEEDLPDIFATSVIVFLYISLAL	651			
248		WQRTYVEVHVHQSVAQNSTQKVLSTFT--TTLDILKSFSDSVIRVASGYLLMLAYACLTM	306			
652		GSYSSSRVMVDSKATILGGVAVLCAVMAAGFFSYLGIRSLVILQVPPVLVSVGA	711			
307		---LRMD--CSKSQGAAGVAGVLLVALSVAAAGLGLCSLIGISFNAATTQVLPFLAGVG	361			
712		DNIFIFVLEYQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSBAICFFLGAALTPMPAVRT	771			
362		DDVFLLAHASETGQNKRIPEEDRTGECLEKRTGASVALTSISNVTAFFMAALIPALRA	421			
772		FALTSGIATLDFLLQMSAFVALLSDSKQKQASRLDVCC-----VKQEL--	818			
422		FSLQAAVVVNFAMVLLIFAILMDLYRREDRLDIFCFTSPCVSRVIOVEQAYTD	481			
819		-----PPGGGEGI-----	827			
482		THONTRYSPPPYSSSHSFAHETQITMSTVOLRTEYDPHTHYVTTAEPSEISVQPVTV	541			
828		-----LLGPFQKAYAPFLHITRGV	848			
542		QDTLSQSPBSTSSTRDLSQFSDSLHCLPEPCTKWTLSSFAEKHYAPFLKPKAKVV	601			
849		VLLFLALFGVSLYSMCHISVGLDQELAKPDSVLLDYFLFNRYFEVGAPEVFTVLGY	908			
602		VIFLFLGLGVSLYGTTRVRDGLDITDIPRETREYDFIAAQFKYFSF-----Y	650			
909		NFSSEAGNVALCSSAGCNNSFTKIQYAT-----EFPEQSYLAIAPASS-----W	953			
651		NM-----YIVTQADYPNIQHLLYDLHRSFSNVKTVNLEENKQLPKMW	693			
Query Match						
Best Local Similarity						
Matches						
341		FOGWGTVASWPLTILVLSVIVPVALAAGLVFTELTTDPVELWSPANSQARSEK-AFH--	397			
14		FRQLGFLICDHPPLPPFPVFLFTAAAGVGLLHNLPLSDAVLYFTPLGAQSKMERMS	73			
398		-----DQHFGP---FFRTNQVILTAPNRSSVRYDLSLLGPKNFGSGLDLDLLELLEQ	450			
74		WPLTDNNYIPGRAVTSREIQVTALARN---DSNILDPAFANAVQLDKY-----IQTR	124			
451		LRHLQVMSPEAQRNISLQDICYAPLNPDPNTSLDYCCINSLLQYFQNNRTLLLTANOTLM	510			
125		VRVLH-----	129			
511		QOTSQVDMKDHFLYCANAPLTFKQGTALA-----LSCMADYG---APVFPFLAIGGY-	559			
130		-----NGHYYSYKNLCIQYKNGGCPNKHVHILSDLNHGFNITYPYRFGSEGGYI	181			

Qy	560	----	-KGKDYSE--AEALIMTFSLNVP-----AGDPLRAQAQLWEAFLEE	599
Dd	182	GSSGGVTVMGKENETDLASAKAWFMIYHUKFPEEMSYTSGE-----WE---	LBL	230
Qy	600	MRAFQRNAGMF-QVTFTAERSLDEINRTTAEDLIPATPSYIVIFYISIALGSY--SS	656	
Dd	231	GRMLTYQPEDPVISITYPHSQTLADELKRNADTLIPRIISITLLIVPSTLCSLSPIDGS	290	
Qy	657	WSRVWDSKATILGIGGVAVTVGVNAWGPFYSYGIRSSLVLQVPPVLVSUGADNIFI	716	
Dd	291	FSDIWLKSPTLSILGLVVVSAGIALITGVGFSLMGMYPN-DIVGVMPFLVLAVGVDMNFI	349	
Qy	717	FVLEYORLPREPGEPREVH---IGRALGRVAPSMILLCSLSEAICFELGALTMPMAVRPAL	774	
Dd	350	MVAAV----RRTSRTHVHERMGECLEADAASIIITSTDVLSFGVGAIITPIPAQIFCV	403	
Qy	775	TSGLAVIDFLQMSAFVALLSDBSKRQEASR-----LDVCCCKVK-----	814	
Dd	406	YTGVAIFFAFTYQITTFPAACLALAMKEASGENSLFLIEAVSAEKTSLSFTQRLFNLS	465	
Qy	815	-POBELPPPQGEGELLGFPQKAYAPPFLHWITRGVWLLF-LALFVGLYSKMCHISVGLD	872	
Dd	466	VPDHSASHDVKPQLTSTRPFGEWYAPVLMHPVPGIAMVMWFVIYELGAS-YGCSRKEGLE	524	
Qy	873	QELALPKDSYLLDYFLFNRYF-EVGAPVYFV-----TTLGYNFSSE	913	
Dd	525	PVNLLVEDSVAIPHRYLLEKYFKYGOQVOIVINNAPDLRNHTSRDRVHAMVLDFAFSKH	584	
Qy	914	A-GMNAICSSAGCNNFSTOKIQIYATBFPEQ-----SYLAIPASSWVDVDFDWL	961	
Dd	585	AIGMESV-----QFWLFEMEYYOKELEVQIIDSSFVGLLHHFLASKTNPNPLAEDIYW-	637	
Qy	962	TPSSCCLHYTSGPNKDFPCPSTVNSLNCLKNCMSITWGSVRPSVEQFHLYLPWFINDRPN	1021	
Dd	638	-----GPMPD-----	642	
Qy	1022	IKCPKGLAAVYSTVNLTSDGOVLAS-RFMAYHKPLKNSODYTEALRAARELANITYADL	1080	
Dd	643	-----DDNGTWKVSFRFILGMKDVLVTMDQTDTATWSFREVARW----	681	
Qy	1081	RKVPGTDPAPEVFPTYITNVFEYOYLIPBGLFMLSCLVPTPAVSCLLGLDLRSGLL	1140	
Dd	682	-----PEENVITFMFINMTDQYIIIPNTVQNIIIALVMIVIAVLFIPOPMCS-LW	733	
Qy	1141	NLSIVMLVDTVGFMALWDISVNAVSLINLVSAVGMSVBFSVSHITSFAISTKPTWLER	1200	
Dd	734	VALACASIDFGVIGYMTLWGNLDAISMITIIMSIGSFSDVYSAHIAYGYVVSREDTAAGR	793	
Qy	1201	AKEATISWGSAPVAGVAMTNLPGTLVLGLAKAQILOIFFFFNLNLTLLGLHLGLVPLPV	1260	
Dd	794	VKEALSALGWPLSQG-AMSTIIAVSVLADIPAYMI-VTFPFKVVLSTISGLLHGLVPLPV	851	
Qy	1261	ILSVXV-----GPDVNPALALEQKRABEA-----VAAVMVASCPNHPRSVS	1300	
Dd	852	LLSIFVRCGCCIISSPHGHFSAQIEKQIRIAISSSPDLRTVAPURASSPISFPHRLE	911	
Qy	1301	-TADNIYVNSHPEGSIX	1316	
Dd	912	YTDSEPVTWNRSKNSIK	928	

RESULT 42

```

: RESUBMIT 42
: US-10-890-776A-4805
: Sequence 4805, Application US/10890776A
: Publication No. US20050129683A1
: GENERAL INFORMATION:
: APPLICANT: Zhang, Jian
: TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
: FILE REFERENCE: PB0177
: CURRENT APPLICATION NUMBER: US/10/890,776A
: CURRENT FILING DATE: 2004-07-14
: PRIOR APPLICATION NUMBER: US 10/060,756

```

Db	520	-----NEWDEPEN--TP-----YAMGKNSTKFW-----LRDYENP 548
Qy	994	MSITGMSVRPSVEQ-FHKYLPWFLLNDRENIKCPKGG-----AAYSTSVNLTSDQVLA 1047
Db	549	LYSPFISELEDEEENFYDLLEWFLKS-PGFSHWGDLWMDNKTDYETI-----VKKF 600
Qy	1048	RFMAYHKPLKNSQDYTEALARARELANITADLRKVPCTDPAFVFPYTIINVFYEQVLT 1107
Db	601	RFTTGKDLSTWDRLLTKTWGVADEY-----PDFNVTVDEDAFLDQILS 649
Qy	1108	ILP---EGFLMSLCVPTFAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDISYN 1164
Db	650	IGPTTQSIWTWLCM---AVVCELPNPNVTVITVSTASIDIGVGFSLMGVDLD 705
Qy	1165	AVSLINLVSAVMSVEFVSHITRSPAIS--TKPTWLERAKEATISMGSAVFAGVAMTNLP 1223
Db	706	PISMITIISIGSVDFSAHIAHYFVRSHGSETPDERLADALEALGWPIQ--AALSTILC 764
Qy	1224	ILVGLAKAQLIQIFFRNLNLLITLLGLHLGLVFLPVILS 1263
Db	765	VLPLLFVPSYMW-VVFFKTFIFLVVWVIGLHLGLIFLFIILS 803
RESULT 43		
US-10-060-756A-3		
; Sequence 3, Application US/10060756A		
; Publication No. US20030046717A1		
; GENERAL INFORMATION:		
; APPLICANT: Zhang, Jian		
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN		
; FILE REFERENCE: PB0177		
; CURRENT APPLICATION NUMBER: US/10/060,756A		
; CURRENT FILING DATE: 2002-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00667		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00664		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00669		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00665		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00668		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00663		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: US 09/864,761		
; PRIOR FILING DATE: 2001-05-23		
; PRIOR APPLICATION NUMBER: US 60/327,898		
; PRIOR FILING DATE: 2001-10-09		
; NUMBER OF SEQ ID NOS: 4804		
; SOFTWARE: Acomica Sequence Listing Engine		
; SEQ ID NO 3		
; LENGTH: 954		
; TYPE: PRT		
; ORGANISM: Homo sapiens		
US-10-060-756A-3		
Query Match 8.5%; Score 588; DB 4; Length 954;		
Best Local Similarity 21.8%; Pred. No. 2,4e-41;		
Matches 219; Conservative 190; Mismatches 356; Indels 240; Gaps 39;		
Qy	334	HT-----LLGQFFGCGTGWASWPLTILVSVIPVVALAAGLVF--TELTTDPVELWSAP 386
Db	115	HTDCLGLSRTFQWLQGVQGAHPWIFLLAPLMTAALGTGFLYLPKDEEDLEEHTPV 174
Qy	387	NSQARSEKAFHQHGFPGFTTNQVILLAPNRSS-----YRYSLLLGPKNFSGIL 436
Db	175	GSPAKAERRFVQGH-----FTNDSYRFSASRRSTEANFVLLVVSYSDSLDPAFVRS 230
Qy	437	DLDLLELLELLEQLERLRLHQLVMSPEAQRNISLQDIC--YAPLNPDNTSLYDCCI--NSLLQ 492
Db	231	KLDGAVQDLRV-----AREKGSQIQVQVCARYAL-----CVPPNPILY 270
Qy	493	YFQNNRTLLLTANTQILMGQTSQVDMKDFLYCANAPLTFKDGFTALASCMADYGA-PVF 551
Db	271	AWQNKTLNL-----SSISPAY-----NHRGRPLY 296
Qy	552	PFLAIGY-----KGDYSEAEALIMTFSLNYPAGDPRL-AQAKLWEAEFLSEWRAP 603
Db	297	LTGFGGVIYLLGSLGMLQKLLRAKAMRLLYLK---TEDPEYDVQSKQMLTHLLDQFTNI 353
Qy	604	QRMA--GMQVVFYA-ERSLEDEINNTAEDLPFATSYIVIFLYISLALGSSYSSMRV 660
Db	354	KNILALKKIEVWHFVTSRQLEPEATSVTV--IPVFLAYLILILFATVSCFRD----- 406
Qy	661	MVDSKATLGLGGVAVILGAVMAAGFFSYLGIRSLVILQVVPFLVLSVAGDNTPIFVLE 720
Db	407	CIRNMCAVAFGVSFAFLAVVSGFLLHIGV-FPVIIVANSFPLILGVGDDFMISA 465
Qy	721	YORLPFRGPGEPRVHIGRALGRVAPSMLLCSLSAICFFLIGALTPMPAVRTPALTSLAV 780
Db	466	WHK-TNLAGDIRE-RMSNVYSKAAVSITITITNIALYTIGMSFRSVQCFCIYTGTTL 523
Qy	781	ILDFLQMSAFVALLSLDSKQEAERLDVCCVKPQELP-----PPQGG 824
Db	524	LFCFYFNITCFGAFWALDGREV-----VCLWLKADPKWPSKPKCCFPFGSVDEHG 578
Qy	825	EGL--LLGFFQKAYAPFLLHMITRGVLLLFALFGVSLYSKMSCHISVGLDQELALPKDSY 882
Db	579	TDIHPMSLFFRDYFCGFLTRSESQYFVVIIVLYIISIIYGVGFHQEGLDLRLNLASDDSY 638
Qy	883	LLDYFLFLNRYF-EVGAPVYFVTLLGYNFSSEAGNAICSSAGCANNFPTQKIQYATEPP 941
Db	639	ITPVFNVEENYFSDYGPVVMVITVKVDYWDK-----DVRQKLENCIKIF 683
Qy	942	EQSYLAIP--ASSWVDDFDIDMLTPSSCCRLYISGPNKDKFCPSTVNSLNLKNCMSITMG 999
Db	684	EKNVYVDKNLTFELWDAYVQ-----YLGNSQD---PNEKNT----- 717
Qy	1000	SVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSDQVLAASFMAHYHKLKNS 1059
Db	718	-----FMNNIPDFLSNFPN-----FQHDINISSNEIISRGFTQTTDVSSS 759
Qy	1060	QDYTEALRAARELANITADLRKVPCTDPAFVFPYTIINVFYEQVLTLP----- 1111
Db	760	-----AKKI---LLFQLARI-AEDCQIPLVYNOAFIYFDQYAAILEDTVRNVLVA 807
Qy	1112	--GLFWLSLCVLP-----TFAYSCLLGLDLRSGLNLLSIVMILVDTVGFMAWDI 1161
Db	808	SAAMFIVSLLIPLCLSLWVTFALGSVIGV-----TGFWAFKV 848
Qy	1162	SYNAVSLINLVSAVMSVEFVSHITRSPAISATKPTWLERAKEATISMGSAVFAGVAMTNL 1221
Db	849	NLDSISMINLVICTGFSDFSAHISYAFVSSQPSVNQKSVREALYLLGYFVLQS-AISTI 907
Qy	1222	PGILVGLAKAQLIQIFFRNLNLLITLLGLHLGLVFLPVILSVG 1266
Db	908	IGVCVLAARAKAVIFRT-PFKIMFLVMIFGAHGLFIPFVPLTFFG 951
RESULT 44		
US-10-890-776A-3		
; Sequence 3, Application US/10890776A		
; Publication No. US20050129683A1		
; GENERAL INFORMATION:		
; APPLICANT: Zhang, Jian		
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN		
; FILE REFERENCE: PB0177		
; CURRENT APPLICATION NUMBER: US/10/890,776A		
; CURRENT FILING DATE: 2004-07-14		
; PRIOR APPLICATION NUMBER: US 10/060,756		
; PRIOR FILING DATE: 2002-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00663		
; PRIOR FILING DATE: 2001-01-30		
; PRIOR APPLICATION NUMBER: PCT/US01/00664		
; PRIOR FILING DATE: 2001-01-30		

Db	403	LLMLAYACVTN---LRWD---CAQSGSVGLAGVLVVALAVASGLGLCALLGITFNAAATQ	457
Qy	701	VVPELVLSVGNADNIFIVLEYQRLPRPGEPREVIHIGRALGRVAPSMCLCSLSAICFFL	760
Db	458	VLPELALGIGVDDVFL--LAHAFTEALPGTQLQERMGECLQRTGTSVVLTSINNMAAFM	515
Qy	761	GALTMPAVRTALTSLGLAVILDFLQMSAFVALLSDSKQKQASRLDVCC--	812
Db	516	AALVPAPALRAFSLQAAIIVWGCTFVAVMLVFPAILSLDLRRRHQRDLVLCFSSPCSAQ	575
Qy	813	---VKPQEL-----	827
Db	576	VIQILPQELGSGTVPVGIHAHTATVQAFTHCEASSQHVVTILPQAHLVPPSPDPLSEL	635
Qy	828	---LLG-----	855
Db	636	FSPGSTRDLLGOEBETROKACKSLPCARNWLAHFARYQAPALLQSHAKAIVLVFGA	695
Qy	856	LFVGLSYSMCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAQVYVTTTLYGNFS-SEA	914
Db	696	LLGLSLYGATLVQDGLALTVDVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHQR	754
Qy	915	GMAIACSSAGCNFSFTQKIQYATEFPQSILAIPAS-----SWVDDEIDMLTPSSCC---	967
Db	755	AL-----FDLHORFSSL-----KAVLPPPATQAPRTWLHYRNWLOGIQAAFDQ	798
Qy	968	---RLYISGNKDKFCPSTVNSLNCIKNCMSITMGSRVPSVQPHKLPWFPLND	1018
Db	799	DWASGRITHSYRNG-SED-----GALAYKLLIQTGDAQRPDL-FSLQTRKLVD	846
Qy	1019	RNIKCPKGLAAYSTSVNLTSDGQVLASRFMAYHKP-----	1055
Db	847	REGLIPPE--LFYMGTLTVWVSDPLGLAASQANFYPPPEWLHDKYDTTGENLRIPPAQP	904
Qy	1056	---LKNQSDYTEALRAARELAANI-TADLRKVPGTDPAPFVPPYTTITNVFY	1102
Db	905	LBFAQFPFLRLGKTADFEAIEGARACACBAGQVGHAYPSGPF-----LFW	954
Qy	1103	EQVLTLPBGLFMLSIC--LVPTFAVSCILLGLDLRLSGLNLLSIVMLVDTVGFMALWD	1160
Db	955	EQVGLG--RRCELLAVCILVCTFLVCALLLNPNWTAGLI-VLVLAMTVLFGIMGFLG	1011
Qy	1161	ISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVFAVAMTN	1220
Db	1012	IKLSAIPVVLVASVIGVGFVHVVALGF-LITQGSRLRAAAHLEHTFAPVTDG-AIST	1069
Qy	1221	LPGLVLGLAKAQLIOIFPRLNLLITLGLLHGLVFLPVLISVYVGPVNPALALEQKRA	1280
Db	1070	LLGLMLAGSHDFIVRYFAALTUTLGLLHGLVLLPVLISILGPP--PEVIQMYKES	1127
Qy	1281	EEAFAAV-----MVASCNPHPRSVSTADNIYVN	1308
Db	1128	PEILSPAPQGGGLRWGASSLSPQSFARVTTSMTVAIH	1165
RESULT 46			
US-09-990-046-7			
; Sequence 7, Application US/09990046			
; Patent No. US20020156245A1			
; GENERAL INFORMATION:			
; APPLICANT: de Sauvage, Frederic			
; APPLICANT: Carpenter, David A.			
; TITLE OF INVENTION: Patched-2			
; FILE REFERENCE: P1405R1			
; CURRENT APPLICATION NUMBER: US/09/990,046			
; CURRENT FILING DATE: 2001-11-20			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/293,505			
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-15			
; NUMBER OF SEQ ID NOS: 32			
; SEQ ID NO 7			
; LENGTH: 1182			
; TYPE: PRT			
; ORGANISM: Mus musculus			
US-09-990-046-7			
Query Match			
Best Local Similarity 8.4%; Score 583.5; DB 3; Length 1182;			
Matches 282; Conservative 154; Mismatches 433; Indels 297; Gaps 44;			
Qy	330	SPSTHTLLGQ-----PFQG-----WGTWASWPLTILVLSVIPVVALAAGLVITELT	376
Db	20	SSAPHILAGSLQAPLWLRAYFQGLLPSLGCRIQKHCQKVLFLGIVAFGALALGLRVAVIE	79
Qy	377	TDPVELMSAPNSQARSEKAFHDQHFQ-PFRTNQVILITAPNRSYRYVDSLLGPKNFSGI	435
Db	80	TDLEQLWVEGSRVSQSELHYTKELGEEAAYTSQMLI-----QTAHQGGNVLTFE----	131
Qy	436	LDLDLLELELQERLRLQVNSPEAQRNISLODICY-----APLNPDN-----TSLYDCC	486
Db	132	LD-----LHLQAALTASKVQVSLYKGSWLNKICYKSGVPLIENGMIERMIKLPFCV	184
Qy	487	INSLIQVFQNNRTLLLTANQTLMGQTSQVDWKD---HFLYCANAPLT-----FKDGTALA	539
Db	185	ILTPDLCFWEAG---LOGGSAYILPGRPDQIWTNLDQQLLEELGPFASLGGFRELLDKA	241
Qy	540	LSCMADYCAPVF-----PFLA---IGSYKGDYKYS-----	565
Db	242	QVGQAYVGRPCLDPPDCHPCPSAPNRHSRQAPNVAQELSGGCHGFSHKFMHWBELLGG	301
Qy	566	-----EAEALIMTF-----SLNNYPAGDPRLAAQAKLWEEAFLEEMRAFORRMAG	609
Db	302	TARDLOQLLRAEALQSTFLLMSPRQLYEHFRGDIQTHDIGNWSEEQASWVQAQWRRFVQ	361
Qy	610	MFQVTFTAERLE-----DEINRTTADLPFATSIYIV--IFLYISIALGSIYSWS	658
Db	362	LAQALPANASQOIHAHPSSTLDDILRAFSE-----VSTTRVVGYYLLMLAYACVTMLRWD	417
Qy	659	RMVDSKATLGLGGVAVVLGAVMAAMGFSSYLGRSSLVIIQVPPFLVLSVGADNIFIV	718
Db	418	--CAQSGAVGLAGVLLVALAVASGLGICALLGITFNAATTQVLPFLGALGIVDDIFLLA	475
Qy	719	LEYQRLPRRPEPREVHIGRALGRVAPSMCLCSLSEACFFLGLALTMPAVRTPALTSGL	778
Db	476	HAFTKAP--PDTPLPERMGECLRTGTSTVALTSVNNMVAFMAALVPIPALRAFSLQAAI	533
Qy	779	AVILDFLLQMSAFVALLSLSKQASRLDVCC-----VKPQE-----	817
Db	534	VVGCNFAAVMLVFPAILSLDLRRHRQLDVLCCFSSPCSAQVIQMLPQELGDRAPVGI	593
Qy	818	-----LPP-----PG-----QGEQ--	826
Db	594	AHLTATVQAFTHCEASSQHVVTILPQAHLLSPASDPLGSELYSPGSGSTRDILLSQEBGTG	653
Qy	827	-----LLLGFFQKAYAPFLHWHITRGVVLLFLALFGVSLYSMSCHISVGLDQ	873
Db	654	POACRPLLCAHWTLAHFARYQFAPLLIQTTRAKALVLLFFGALLGLSLYGATLVQDGLAL	713
Qy	874	ELALPKDSYLLDYFLFNRYFEVGAQVYVTTTLYGNFS--SEAGMNAICSSAGCNPSFTQ	932
Db	714	TDVPRGTKEHAFLSAQLRYFSL-YEVALVTQGGFDYAHQSQRAL-----FDLHQ	761
Qy	933	KIYATEPPEQSYLAIPAS-----SWDDF-----IDWLTSPSSCCRLYISGPNK	976
Db	762	RFSSL-----KAVLPPPATQAPRTWLHYRWSWLOGIQAAFDQDQWASGRITCHSYRNGSD	816
Qy	977	D-----KPCPSTVNSLNCIKNCMSITM-----GSRVPSVEQFHKLPWFLNDRPNIKCPK	1026
Db	817	GALAYKLLIQTGNAOEPLDFSQITTRKVDKGLIIP--ELFYMGITVWVSSDPL-----	869
Qy	1027	GGIAA-----YST-----SYNLTSDGQVLASRFMAYHKPKNSQDYTEALR	1067
Db	870	GLAASQANFYPPPEWLHDKYDTTGENLRIPAAQPLEFAQFPFLHGLGKTADFAVEAIE	928
Qy	1068	AARELAANI-TADLRKVPGTDDPAFEVPPYTTITNVFYEQVLTILPEGLFMLSIC--LVPTF	1124

Db 929 GARAACTEQAGVHAYPSGSPF-----LFWQVQLGL--RRCFLLAVCIILLVCTP 976
QY 1125 AVSCLLGLDLRSGLNLLSYIMLVDTVGFMALMDISYNAVSLINLNSAVGMSVEFVSH 1184
Db 977 LVCALLLLSPWTAGLI-VLVLAMTVLFGIMFLGKLSAIPVVLVASIGVFEVTH 1035
QY 1185 IYRSPAIATKPTWLERAKATISMSANVAGY---AMTWLPGILVLGLAKAQLIQIPPR 1241
Db 1036 VALGFLTSHSGSNLRAA-----SALEQTPAPVTDGAVSTLLGLLMLAGSNFDFIIRYFV 1090
QY 1242 LNLITLTLGLLGLPLPVLVSVP 1267
Db 1091 VLTVTLGLLGLLPLVILLSILGP 1116

RESULT 47
US-10-425-115-328081
; Sequence 328081, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: Plants
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 328081
; LENGTH: 343
; TYPE: PRN
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_62274C.1.pap
US-10-425-115-328081

Query Match 8.4%; Score 582.5; DB 4; Length 343;
Best Local Similarity 35.8%; Pred. No. 1.5e-41;
Matches 126; Conservative 71; Mismatches 116; Indels 39; Gaps 10;

QY 848 VVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTLG 907
Db 7 VVITFVGLTFASIALSTRQLGLEQIVLPNSYLQDYFDLAKYMKVGPPLYPVWK-D 65
QY 908 YNFS--EAGMNAICSGAGNPFQTKIQYATEPEEQSVLAIPASSWDDFDNLTTPSS- 965
Db 66 FNYSSAVDTNQCISQCNLSNLSLEISRQSLSPETSIIAKPAASWLDLFLNMSPEAF 125
QY 966 -CCRLYISGPNKDKFCP-----STVNSLNLKNCMSITWGS----VRPSV 1005
Db 126 GCCKRFVNG-----SYCPDDQPPCQLDQVSGSCMTKTC-SNCTTCFLHSLDNGRST 180
QY 1006 EQFKYLPWFLNDRENIKCPKGLAAYSTVNL--SDGQVLAIRFMAHYKPLKNSQDYT 1063
Db 181 TQFRDKLPWFLDALFSSDCSGKGAYSTSLDLSVSGEIIQASAFRTVHTPLNKSQDYV 240
QY 1064 EALRAARELANITADLRKVPDTPAFVPPYITNVFEQYLTILPEGLFMLSCLVPT 1123
Db 241 NSMARARDFSSKMSRDLO-----MKIFFYSVFYIPFEQYLSVMKTAIWNICVCLGTI 292
QY 1124 PAVSCLLGLDLRSGLNLLSYIMLVDTVGFMALMDISYNAVSLINLNSAV 1175
Db 293 FVV-CFIVTSSLWASAILTLVLAIVLMDMGVMAILGILQNLNATSVVNLVMSI 343

RESULT 48

US-09-909-280A-2
; Sequence 2, Application US/09909280A
; Patent No. US20020160375A1
; GENERAL INFORMATION:
; APPLICANT: Buncroft, David A.

; TITLE OF INVENTION: HUMAN PATCHED GENES AND PROTEINS, AND USES RELATED
; TITLE OF INVENTION: THERETO
; FILE REFERENCE: CIBT-P02-050
; CURRENT APPLICATION NUMBER: US/09/909,280A
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 09/207,857
; PRIOR FILING DATE: 1998-12-08
; PRIOR APPLICATION NUMBER: US 60/067,940
; PRIOR FILING DATE: 1997-12-08
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1203
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-909-280A-2

Query Match 8.4%; Score 580; DB 3; Length 1203;
Best Local Similarity 22.9%; Pred. No. 1.8e-40;
Matches 283; Conservative 172; Mismatches 473; Indels 310; Gaps 42;

QY 306 RVAPARDKSKMVDPKKGTSLSDKL--SPSTHTLLGQFFOG-----MGTWVASWPLTILVLS 359
Db 3 RSPPLRELPSPSYTPPARTAAPQILAGSLKAPLWLRAYFQGLLFLSCGCIQRHCGKVLFLG 62
QY 360 VIPVVALAAGLVFTBLTDPVELWSPNSQARSEKAFHDQHFQ--PFFRTNOVILTAPNR 417
Db 63 LLAPGALALGLRMAIETNLEQLWVEVSGSRVSEQLHYTKELGERAAVYSQMLIQARQE 122
QY 418 SSYRYSLLGPNFSGILDLDLLELLELQRLHQLRHLQVWSPQAQRNLSLDICY---AP 474
Db 123 GEN-----LTPALG-----LHLQAALTASKVQVSLYKGSWDLNLCYKSGVP 166
QY 475 LNPDN-----TSLYDCCINSILQVFNQNRITLLLTANQTLMGQTSQVDWMD- 520
Db 167 LIENGMIERMEIKLPFCVILTPDLCFWGAK---LQGSAYLPGRPDIIQWTLNDPEQLLE 223
QY 521 -----HFLYCANAPITPKDGTALALSCMADYGA 548
Db 224 ELGPPASLEGPRELLDKAQVQVYGRPCLHPDDLHCPPSAPNHHSRQAPNVAHELSGGC 283
QY 549 PVPPP-----LAIGYKGDYSE---AEALIMTP-----SLNYPAGDPRLAAQKL 591
Db 284 HGFSHKPMHOBELLGGMWARDPOGELLRAEALQSTFLMSPRQLYHEFRGDTYQTHDIGW 343
QY 592 WEEAFLBEMRAFQREBWMAGFQVTPFAERSLE--DEINRTTAEDLPFAPTS-----Y 640
Db 344 SEEQASTVLAQWRFPVLAQALPENASQIIHAFSSTLDDI-LHAFSEVSAARVVG 402
QY 641 IVIFLYISLALGSYSWSRVVDKATLGLGVAVVLGAVMAAGFFSVLGRSSILVILQ 700
Db 403 LLMLAYACVTM---LRWD--CAQSQSVGLAGVLVALAVASGLGICALLGITFNAAV 457
QY 701 VVPLVLVSGADNIPFVLEYORLPRRPGEPREHIGRALGRVAPSMILCSLSEACPL 760
Db 458 VLPLGIGVDDVFL--LAHAFTALPGTPLQRMGECLQRTGTSVVLTSINNMAAPLM 515
QY 761 GALTMPAVRFPALTSGLAVIDLFLLOMSAFVALLSLDSKQESRLDCCC-----812
Db 516 AALVFPALRAFSLQAAIIVGCTFVAVMLVFPATFSLDLRRHRCORLDVLCFSSPCSQA 575
QY 813 ---VKPQEL-----PPKQGBGL-- 827
Db 576 VIQILPQELGDGTVPVGIHAHLTATVQATHCESASHVVTILPQAHVLPFPSPDPLGSEL 635
QY 828 -----LIG-----PFQKAYAPFLLHMTITRGVILLFLA 855
Db 636 FSPGSTRDLLQGBETRQKAACKSLPCARWNLAHFARYQAPAPLQLQSHAKAIVLVFCA 695
QY 856 LFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTLGYNPS-SEA 914
Db 696 LLGLSLGATLVQDGLALTVDVVRPGTKBHFASLAQLRYFSL--YEVALVTQGGFDVAHSOR 754

QY 915 GMAICSSAGCNFSFTQKIYATEFPQSYLAIPAS-----SWDDFDIDWLTSPSSCC--- 967
DB 755 AL-----FDLHQRFSSL-----KAVLPPPTAQPRWLHYRNWLGQQAARDQ 798
QY 968 -----RLYISGPNKDKFCPSTVNSLNCINCKMCSITMGSVRSVQFHKYLPFWFLND 1018
DB 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQEPFLD-FSQLTTRKLVLD 846
QY 1019 RPNKICPKGGLAAYSTSVNLSDGOVLASRFMAYHKP----- 1055
DB 847 REGLIPPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEWHLHDKYDTTGENLRIPPAQP 904
QY 1056 -----LKNQSDYTEALRAARELAANI-TADLRKVPETDPAFEVFPYTTINVPY 1102
DB 905 LEFAQPFPFLRLGLOKTADFEAIEGARAACAEAGQAGVHAYPSGSPF-----LFW 954
QY 1103 EGYLTILPEGLPMLSIC--LVPTFAVSCILLGLDLRSGLNLLSIVMLVDTVGFPMALWD 1160
DB 955 EGYLGL--RRCPFLAVCILLVCTFLVCALLLNPNMAGLI-VLVLAAMTVLVELFGIMGFLG 1011
QY 1161 ISYNAVSLINLVASVGMSEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
DB 1012 IKLSAIPVIVLVASVIGVEFTVHVALGF-LTQGSRLRAHALEHTFAPVTDG-AIST 1069
QY 1221 LPGILVLGLAKAQLIQIPFERLNLLITLGLLHGLVFLPVLVSIVYVGPVDPNPALEQKRA 1280
DB 1070 LUGLLMAGSHDFVIRVYFFAALTITLGLLHGLVFLPVLVSILGPP--PEVIQMYKES 1127
QY 1281 BEAFAV-----MVASCPNHPFSRVSTADNIYVN 1308
DB 1128 PEILSPAPQGGGLRWGASSLPQSFARVTTSMVTVAIH 1165

RESULT 49
US-10-791-844-1
; Sequence 1, Application US/10791844
; Publication No. US20040146929A1
; GENERAL INFORMATION:
; APPLICANT: ZAPHIROPOULOS, Peter et al.
; TITLE OF INVENTION: A NOVEL COMPONENT IN THE HEDGEHOG SIGNALLING PATHWAY
; FILE REFERENCE: 2921-0130P
; CURRENT APPLICATION NUMBER: US/10/791,844
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: US/09/807,007
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-791-844-1

Query Match 8.4%; Score 579; DB 4; Length 1146;
Best Local Similarity 23.3%; Pred. No. 2e-40;
Matches 27%; Conservative 161; Mismatches 452; Indels 298; Gaps 40;

QY 306 RVAPARDKSMVDPKKGTSLSKL--SFSTHTLLGQFFQG-----WGTWASWPLTILVLS 359
DB 3 RSPPLRELPPSVTPPARTAAPQILAGSLKAPLWRAVFOGLLSLGCIGIQRHCGKVLFLG 62
QY 360 VIPVVALAAGLVPELTTPDELWASPNQARSEKAFHDOHFG--PFEETNOVILTAPNR 417
DB 63 LIAFGALALGLRMAIETLEQLNEQVGVSEVGSQELHTEKUGEEAAYSQMLIQTAQRE 122
QY 418 SSRYSDSLGLGPNFSGILDLDLLELLELQERLRHLQVSPQAQRNLSLDICY---AP 474
DB 123 GEN-----ILTPEALG-----LHLQAALTASKQVSLYSGSWDLNKKCYKSGVP 166
QY 475 LNPDN-----TSLYDCCINSLLOYFQNNRTLLLTANQTLMGQTSQVDWKD----- 520
DB 167 LIENGMIERMIKLPFCVILTPIDCFWEGAK---LOGGSAYLPGRPDIQWTLNLDPQLLE 223

QY 521 -----HFLYCANAPLTFKDGKTALALSCMADYGA 548
DB 224 ELGFPASLEGPRELLDKRAQVQAVGRPCLHPDDLHCPPSAPNHSRQAPNVVAHELSCG 283
QY 549 PVFPF-----LAIGYKGYKQYSE--ABALIMTP-----SLNYPADPRLAQAQKL 591
DB 284 HGFSHKFMHMQEELLGGMARDPOGELLRAELAQSTFLLMSPRQLYEHFRGDIYQTHDIGW 343
QY 592 WEAPLEBEMAFORMAGMFQVTTAERSLE-DEINRTAEDLPIFATS-----Y 640
DB 344 SEQASTVLOAQWRRFVQLAQEALPENASQIHAFFSSTLDDI-LHAFEVSAARVGGY 402
QY 641 IVIFLYISLALGSSYSSRVNVDSKATILGLGGVAVVLGAVNMAAGFFSYLGSIRSLVILQ 700
DB 403 LLMLAYACVTM--LRWD--CAQSQGSVGLAGVLLVALAVASGLGCLALGITFNAATQ 457
QY 701 VVPLFIVSVGADNIPFVLEYQRLPRRPGEBREVIHGRALGRVAPSMILCSLSEALCFPL 760
DB 458 VLPFLALIGVDDVFL--LAHAFTALPGTFLQERMGECLQRTGTSVVLTSSINNMAAFILM 515
QY 761 GALTMPAVRTFALTGLAVILDFLLQMSAFVALLSLDSKQASRLDVCCC----- 812
DB 516 AALVPIPALRAFSLQAAIIVGCTFVAVMLVFPAILSLDLRRRHQCORLDVLCFSSPCSAQ 575
QY 813 ---VKPQEL-----PDLHQRFSSL-----PPGQSGEL-- 827
DB 576 VIQILPQELGDTVPVGGIAHLTATVQAFTHCEASSQHVVITILPQOHLVPPSPDPLGSEL 635
QY 828 -----LLG-----FFOKAYAPFLLHMITRGVLLFLA 855
DB 636 FSPGSTRDLLGQBEETRQKAACKSLPCARWNLAHFARYQFAPLLLOSHAKAIVLVLFGA 695
QY 856 LFGVSLYSMCHISVGLDQELALPKDSYLLDFLNRVFEVGAQVVFVTTILGYNS-SEA 914
DB 696 LGLSLYGATLVQDGLALTVDVPRGTKEHAFSLAQRLYFSL-YEVALVTQGGFDVAHSOR 754
QY 915 GMAICSSAGCNFSFTQKIYATEFPQSYLAIPAS-----SWDDFDIDWLTSPSSCC--- 967
DB 755 AL-----FDLHQRFSSL-----KAVLPPPTAQPRWLHYRNWLGQQAARDQ 798
QY 968 -----RLYISGPNKDKFCPSTVNSLNCINCKMCSITMGSVRSVQFHKYLPFWFLND 1018
DB 799 DWASGRITRHSYRNG-SED-----GALAYKLLIQTGDAQELLD-FSQLTTRKLVLD 846
QY 1019 RPNKICPKGGLAAYSTSVNLSDGOVLASRFMAYHKP----- 1055
DB 847 REGLIPPE--LFYMGILTVMVSSDPLGLAASQANFYPPPEWHLHDKYDTTGENFRIPPAQP 904
QY 1056 -----LKNQSDYTEALRAARELAANI-TADLRKVPETDPAFEVFPYTTINVPY 1102
DB 905 LEFAQPFPFLRLGLOKTADFEAIEGARAACAEAGQAGVHAYPSGSPF-----LFW 954
QY 1103 EGYLTILPEGLPMLSIC--LVPTFAVSCILLGLDLRSGLNLLSIVMLVDTVGFPMALWD 1160
DB 955 EGYLGL--RRCPFLAVCILLVCTFLVCALLLNPNMAGLI-VLVLAAMTVLVELFGIMGFLG 1011
QY 1161 ISYNAVSLINLVASVGMSEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTN 1220
DB 1012 IKLSAIPVIVLVASVIGVEFTVHVALGF-LTQGSRLRAHALEHTFAPVTDG-AIST 1069
QY 1221 LPGILVLGLAKAQLIQIPFERLNLLITLGLLHGLVFLPVLVSIVYVGP 1267
DB 1070 LUGLLMAGSHDFVIRVYFFAALTITLGLLHGLVFLPVLVSILGPP--PEVIQMYKES 1127

RESULT 50

US-10-890-776A-4807
; Sequence 4807, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177


```

; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 4807
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Sequence of the Patched motif of gi 1825729
US-10-890-776A-4807

Query Match      8.1%; Score 563; DB 5; Length 836;
Best Local Similarity 23.2%; Pred. No. 3e-39;
Matches 229; Conservative 162; Mismatches 314; Indels 282; Gaps 34;

QY 377 TDVPLMSAPNSQARSEK-AFH-----DQHFGP----PRTNQVILTAPNRSYRYDSSL 426
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
4 SDAYILFTPLGAKQMRMSIEKMPITDNNIPGRAVTSREIQVTALARN-----DSNI 59

QY 427 LGPKFSGILDLDLLELLEQLERLRLQVMSPEAQRNISLDICVAPLNPNTSLYDCC 486
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
60 LDPKANAVYQLDKY-----IQTRVRLH----- 83

QY 487 INSLQYFONNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLFFKQGTALA-----L 540
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
84 -----NGHYISYKNCLOQYKNGCGPSNKHVHL 111

QY 541 SCWADYG-----APVPEFLAIGY-----KGDYSE-----AEALIMTFSLNYP- 580
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
112 SLDLHNHGFNITYPYFRFGSEGGYIGSSLGVTVMKGENETDILASAKAWFMYHLKFHPE 171

QY 581 -----AGDPLQAQKLWEAEFLMEKRAFORMAGMP-OVTFTAERSLEDEINRTTAEDLP 634
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
172 ENSYISGE-----WE-----LELGRMLTOYEDDPYISITYFHSQTLADELKNADTLIP 220

QY 635 IPATSYIVIFLYISALGSY--SSMSRVMDSKATLGLGGVAVVLGAVMAAMGFFSYLGI 692
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
221 RPIISITLLIVESTLCSLFDGSDWLSKPILSILGVVSAGIAITLGVGFLSLMG 280

QY 693 RSSLVLQVVPVLVLSVGDNFIIFVLEQRLPRPGEPREVH--IGRALGRVARSMLLC 750
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
281 PYN-DIVGMVPEFLVAVGVDNFMFLAAV-----RTSRTHVTHERMGECLADAASVILIT 335

QY 751 SLSEAI CFFPLGALTMPAVRTEALTSGLAVIDLQLQMSAFVALLSLDSKQREASR---- 806
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
336 SSTDLVSFGVGAITTPAQVFCVTGVAIFPAFYQITFFAACLALAMKHEASGRNSLIF 395

QY 807 -LDVCCVK-----PQLPPPGQGGELLGPFQKAYAPFLIHWITRGVV 849
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
396 LIEAVSAEKKTSLSLTFQRLFNLSGVPDHSASHDVKQPLTSRFFGEWYAPVLMHPVVRGIA 455

QY 850 LLLP-LALFGVSLYMCHSISVGLDQELALPKDSYLLDYFLFLNRYV-EVGAPVYV- 903
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
456 MWVFVYILLGAS-YGCSRIKEGLEPWNLLVEDSYAIPHRYLLEKYFMWYQQQVQVINNA 514

QY 904 -----TTLGYNFSSEA-GMNAICSSAGCANNFSFTQKIQYATEPPEQ----- 943
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
515 PDLNHTSRDRVHAMVLDFAATSKHAGMESV-----QFWLEMEERYKQLEVLIDSS 568

QY 944 -----SYLAIPASSWDDFDIWLTPSSCCRLYISGPNKDKFCPSTVNSLNCIKNCMSIT 937
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
569 FYGLLHHFLASKTNPLAEDIYW-----GMPD----- 596

QY 998 MGSVRPSVEQFHKYLFWFLNDRPNIKCPKGLAAVSTSVNLTSDGQVLAS-REWAYHKPL 1056
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
597 -----DNGTMTVKSFRFLGKMDL 615

QY 1057 KNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYTTITNVFYEQLTTLPEGLFML 1116
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
616 VTTMDQTDATMSFREVAARW-----PEFNVTTMPIMWTFDQIIIIIPNTVQNI 664

QY 1117 SLCLVPTFAVSCILLGLDLRSGLLMLLSIVMLVDTVGPMALWDISYNAVSLINLVSAGV 1176
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
665 IIALLVMIIVIAVLFIQPMCS-LWVALACASIDFGVGYMTLWGLVNLDAISMITIISIG 723

QY 1177 MSVEFVSHITRSPFAISTKPTWLERAKEATISGSAVPAGVAMTNLPGIILVIGLAKAQLIQ 1236
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
724 FSDYSAHIAIYGYVRSREDTAAAGRVKEALSALGWPLSQG-AMSTIIAVSVLADIPAYMI- 781

QY 1237 IPFFRLNLIIITLLGLLHGLVFLPVLIS 1263
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
782 VTFFKTVVLSISGLLHGLVFLPVLIS 808

RESULT 51
US-10-890-776A-4806
; Sequence 4806, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 4806
; LENGTH: 793
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Sequence of the HTPL-L Patched motif
US-10-890-776A-4806

Query Match      8.1%; Score 559.5; DB 5; Length 793;
Best Local Similarity 21.6%; Pred. No. 5.6e-39;
Matches 206; Conservative 182; Mismatches 333; Indels 233; Gaps 37;

QY 378 DPVELMSAPNSQARSEKAFHQHGFPPRTNQVILTAPNRS-----YRDSLL 427
```

```
Db      5 DLEHYTPVGSAPKAERPVQGH-----FTTNDSTYRFSASRRSTEFNVSLVSVSDSLL 60
Qy      428 GPKNFSGLDLDLLELELOERLRLQVMSPEAQRNLSLQDIC--YAPLNPNTSLYDC 485
Db      61 DPATAEVSGLDGAQDLRV-----AREKSGQIQYQVCARYAL----- 100
Qy      486 CI--NSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCNAPLTFKGTALALSCM 543
Db      101 CVPPNPILYAWQVNTLNL-----SSISFPAY----- 127
Qy      544 ADYGA-PVPPFLAIGY-----KGKDYSEAEALIMTFSLNNYPAGDPRLL-AQAKLWEE 594
Db      128 -NHRHPLTYTFFGGYILGSLGMLLRAKARLLYLK---TEPEVDVQSKQMLT 183
Qy      595 AFLEEMRAPQRMA--GMQVFTTA-ERSLEDEINRTTAEDLPFATSYIVIFLYISAL 651
Db      184 HLLDQFTNIKILAKKIEVHTSLSRQLEFATSVTV--IPVFHLAYILLPAVTSK 241
Qy      652 GSYSMRSRVMSKATLGLGGVAVLGAAMAAGFFSYLGRSSLVILQVVPFLVLSGA 711
Db      242 PRFD-----CIRNKMCAAFGVISAPLAVVSGFLLHIGV-PFVIIIVANSFPLILGVG 295
Qy      712 DNIFFVLEYQRLPRPGPREVHIGRALGRVAPSMILCSLSEACFFLIGALTTPPAVRT 771
Db      296 DDMFIMISAWHK-TNLADIRE-RMENVYKKAASITITITINILALYTGIMSFRSQVC 353
Qy      772 FALTSGLAVIDLFLQMSAFVALLSLDSKRQBSRLDVCCCVKQBELP----- 819
Db      354 FCITYGTILLFCYFYNITCFGAPMALDGRV-----VCLWLKKADPKWPKFKCCFP 408
Qy      820 ----PQOQEGEL--LLGFQKAYAPFLHWTGRGVLLFLFLALFGVLSYKMSCHISVGLDQ 873
Db      409 FGSVPDEHGTDIHPMSLFRDYGPELTSSESKYFVVFYVLYIISYICGFHVQEGDL 468
Qy      874 ELALPKDSYLLDYFLFLNYF-EVGAPVYFVTLGYNFSSESAGMNAICSAGCNPSPFTQ 932
Db      469 RNLASDVSITPYFNVENYFSDYGRVMVITVKKYVDWK-----DVRQ 513
Qy      933 KIQYATEFPEQSYLAIP--ASSWVDYFIDWLTPSSCCRLYIGPNKXKFCPSTVNSLNC 990
Db      514 KLENTKIFEKVYVDKNLTFFWLDVAVQ-----YLGNSQD---PNEKNT----- 556
Qy      991 KNCMSITMSVRPSVQFHKYLPWFILNDRPNIKCPKGLAAYSTSVNLSDGOVLASRFM 1050
Db      557 -----FMNIPDFLSNPN-----FOHDINISSNEIITSSRGF 589
Qy      1051 AYHKPLKNSQDYTEALRAARELAANITADLRKVGTDPAPFVPPYITITNVFEOYLITLP 1110
Db      590 IQTTDVSS-----AKKKI---LLFQLRRI-AEDCQIPLVYVNAQPIYFDQYAAILE 637
Qy      1111 E-----GLFMLSCLIVP-----TFPVSCLLLGLDLRSLGLNLLSVMTLVD 1152
Db      638 DTVRNVLVASAAMFIVSLILIPYPLCSLWTFPAIGSVIVGV----- 678
Qy      1153 VGFMAUWDSYNVAVNLINVASGVMSVEFVSHITRSPATISTKTWLERAKEATISGSAV 1212
Db      679 TGFMAEFKNLDSISMINLVICTGFSDFPSGAHISYAFVSSSQPSVQKSVEALYLLGYPV 738
Qy      1213 FAGVAMTNLPGILVLGLAKAQLIQIRPPRLNLITLLGLHGLVFLPVLVILSVYG 1266
Db      739 LOS-AISTIIIGVCVLAANKAYIFRT-PFKIMFLVMIFGAAGHLIFIPVLTFFG 790
```

RESULT 52

```
US-10-890-776A-4808
; Sequence 4808, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jjian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
```

```
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acemica Sequence Listing Engine
; SEQ ID NO 4808
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Sequence of the Patched motif of gi 1707052
US-10-890-776A-4808
```

```
Query Match 7.8%; Score 539.5; DB 5; Length 846;
Best Local Similarity 23.0%; Pred. No. 3.4e-37;
Matches 223; Conservative 156; Mismatches 337; Indels 253; Gaps 35;

Qy      401 FGPFRFTNQVIL-----TAPNRSSVYDSLLGPKNFSGILDLLELLE 446
Db      10 FGPYSYTERRIIHDWPLVDGTFVAGRAVTS--REVQAVVAVRSGGILDRVFNELKL 67
Qy      447 LQERLHLQVMSPEAQRNLSLQDICYAPLNPNTSLYDCINSLLYQYFQNNRTLLLTAN 506
Db      68 MESFIRN-NITVQFSNRRTWSFADLCLIA--GPDGR---CANNDHIQ----- 106
Qy      507 QTLMGQTSQVMDKDHFLYCNAPLTFKGTALALSCMADYGAAPVFPFLAIGGYK----- 560
Db      107 --LASRLHQ-----HGINITYPTVRLSDKSAYIAS-----ALGGVKLAKGN 146
Qy      561 GKD-YSEAEALIMTFSLANNYP-----AGDPLAQAKLWEEAFLEEMRAPFORRMAGMFOV 613
Db      147 GENIIVEATAWLLIYQLKFPYNEISVYSG-----LWREFKNQNDY-KQAKYISI 197
Qy      614 TFTAERSLEDEINRTTAEDLPFATSYIVIFLY-----ISLALGSYSSWSRVMDSKAT 667
Db      198 TYFHSQTLSDENLRNAERLAPKFIAGFVILCVFSLCSIVTIKSGSYIDW---VVTKPI 253
Qy      668 LGLGGVAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPRR 727
Db      254 LSVLGVSNAGMGSIASAMGMLTYLEIQYN-DIIAVMPFLVAVGTDMFLMVASLKRTRD- 311
Qy      728 PGEPREHVHIGRALGRVAPSMILCSLSEACFFLIGALTTPPAVRTFALTSGLAVIDLFLQ 787
Db      312 -NLKYDQRIAEACVADAASVILITALTALSGFVGITITIPAVQIFCIYTWCAILLTFAYQ 370
Qy      788 MSFAVALLSLDSKRQBS-----ASRLDV---CCCVKQBELPPFPQOG----- 824
Db      371 LTFEFCALYYVYTRIEBQGLHSIWLRPVATVYSSTPSPLNVKLFWLGSQFQK-PLPSCGVSS 429
Qy      825 -----EGLLGFQKAYAPFLH-WITRGVLLFLFLALFGVLSYLS 863
Db      430 TSSVSTWTSQATSPASKHLHCAATSFRRNWTAPVLMQPMI-RAIAGLWYLYLIGISYIG 488
Qy      864 MCHISVGLDQELALPKDSYLLDYFLFLNR-YFEVGAAPVYFVTT-----LGYNFSSE 913
Db      489 CTHLKEGLEPANLLVDDSYATPHRYVLEKHYWHYGASLIQIVSVNPPDLRDPVERINWDM 548
```

QY 914 AGNNAICSSA-GCNNEP-----TQKQYATE-----PPEQSYLAIPASSWDDFI 958
Db 549 ASTFANCKVAIGDSSVQFLREMQVSEIHKIQYDNEKFDHAAQYIYSDMSQPVVDV 608
QY 959 DMLTPSSCCRLYISGPNKDKPCPSTVNSLNLCKNCSITMGSVRPSVEQFHKYLPWFLND 1018
Db 609 -----WGRNN 613
QY 1019 RPNIKCPKGLAAYSTVNLTSQGLASRFMAYHKPKNSQDYTEALRAARELANITA 1078
Db 614 -----NSERIILKTRFMGLMEDISTTTKQTEATNTPREIASRP-- 651
QY 1079 DLKVPCTDPAEVEFPYITNVPYEQYLILPEGLFMLSCLVPTPAVSCLLGLDLRG 1138
Db 652 -----EQNVITYMPLWLFDTQYALVWNTMODIIVAVACMLVISALLIPQVCSF 702
QY 1139 LNLNLSIVMLVDVTGFMALWDISYNAVSLINLVSAGMSVRFVSHITSFAISTKPTWL 1198
Db 703 WV-AVTIGSIDGLVGLFWLMLNVLDAISMITIISVGFSDYSAHITAYVISKESTTS 761
QY 1199 BRAKEATISMGSVAVFAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITLGLHGLVFL 1258
Db 762 ARVCDALGDLGNPVAQG-AMSTILAVSLSDVPAYMI-VTFKTVPLAISIGFLHGLVFL 819
QY 1259 PVILS-YVG 1266
Db 820 PLMLSVFVG 828

RESULT 53

US-10-415-934-9

; Sequence 9, Application US/10415934

; Publication No. US2004030099A1

; GENERAL INFORMATION:

; APPLICANT: Bayer AG

; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN

; FILE REFERENCE: LIO04 Foreign Countries

; CURRENT APPLICATION NUMBER: US/10/415,934

; CURRENT FILING DATE: 2003-05-05

; PRIOR APPLICATION NUMBER: US 60/245,565

; PRIOR FILING DATE: 2000-11-06

; PRIOR APPLICATION NUMBER: US 60/245,564

; PRIOR FILING DATE: 2000-11-06

; PRIOR APPLICATION NUMBER: US 60/245,572

; PRIOR FILING DATE: 2000-11-06

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 1061

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-10-415-934-9

Query Match

Best Local Similarity 7.7%; Score 531; DB 4; Length 1061;

Matches 185; Conservative 164; Mismatches 322; Indels 180; Gaps 27;

QY 492 QYFQNNRTLLLTANQTLMGQTSQVDWKKDHLFCANAPLTFKDGTAALASCHAD----- 545
Db 5 EVFQELRLDNIIQIATTYDGTYYKDN-----CARNWECENEDILNLDALMDIDIEAQ 61
QY 546 -----YGAPVFPFLAIGYKGYKGYSEAEALIMTFLSNYPAGDPRL----- 586
Db 62 LNLTFPMFNPTWDAHLFP-VFPGTK-----LTEDNYIVSVAIQIVYV 107
QY 587 -----AQAKWSEAFLEENRAFQRMWAGMPQ---VTFTAERSLSEDEINTTADLP 635
Db 108 TADTKRQDAKGAEBETFLRVVGNEN--SGQFKHISVSYPASRTLDHLEKNTKTVPY 165
QY 636 FATSIV--IFYLSIALGYSWSRVMVDSKATGLGCVAVVLGVNVAAMGPFYLGIR 693
Db 166 FSSFTLLMGLFSITCMGD-----AVRSKPFGLMGVNSAIMATLAFLGMYCGI- 217

RESULT 54

US-11-097-143-23451

; Sequence 23451, Application US/11097143

; Publication No. US20050208558A1

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097,143

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

Db 400 TTSGSVSSAYFPYFSTSTLNDILGKSEVSLKNIILGYMFLIYAVTL---IQW-RDP 455
Qy 662 VDSKATILGGLGAVVILGAVMAAGFFSYLGIRSLVILQVPPVLVLSVGADNIPFVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASQTQVPPFLALGLGVQDMFLTHY 515
Qy 722 QRLPRRGEPREVHIGALGRVAPSMCLLSLSEAICFFLGCALTPMPAVRTPALTSGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGSLVSLASLCNMVAFALAAALLPAPAFVFCQAAILL 573
Qy 782 LDFLLQMSAFVALLSLDSKQEARSLDVCCVQPOELPP----- 820
Db 574 FNLGSILLVFPAMISLDRRSARADLLCLLPSPPLPKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGEGL-----LGFQKAYAPPLHMTIRGVVLLFLALFGVSLYSMC 865
Db 634 TROPDLPDVENVTKCCLSVLTWKAKQYAFFIMRPAVKVTSMLALIAVILTSVWGAT 693
Qy 866 HISVGLDQELALPKDSVLLDYPLFNR---YFVGVAPVYFVTLTGYNFSSEAGNNAICSS 922
Db 694 KVRDGLDLTDIVPENT---DEHFLSRQEKYFGF-----YN-----MYAVTQ- 732
Qy 923 AGCWNFSF--TKIQVATEPPEOSYLAIP-----ASSWDDPFIWL----- 961
Db 733 ---GNFEYPTNQKLLY---ETHDQ-FVRIPNIIKNDNGGLTKFWLSLFRDMLLDLQVAFDK 786
Qy 962 -TPSSCCRLYISGNKDKFCPSVTNSLNCIKNCMSITMGSVRPSVEQ----- 1007
Db 787 EVASGCI-----TOEYCKNASDB--GILAYKLMVQTGHVDNPIDKSLITAGHRLVDKD 838
Qy 1008 -----FKHYL-PWFLNDRPNKCPKGGI-----AAYSTSVNLTSDBGVLASRFM 1050
Db 839 GIINPRAFYNSAWATNDALAYGASQGNLKPQORWIHSPEDVHLEIKKSSPLIYTQLP 898
Qy 1051 AVHKPLKNSODYTEALRAARELANITADLRKVPQTDPAFEPVPIYITVFBQYITILP 1110
Db 899 FYLSGLSDYTXSIKTLRSVRDLCLKYEA--KGLPN-----PPSGIPFLFMEQYILYRT 949
Qy 1111 EGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVTGFMALWDISVNAVSLIN 1170
Db 950 SULLALACALAAVP-TAVVILLNNAVAVLTALATVLQLLGVNALLGVKLSANPAVL 1008
Qy 1171 LVSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVIGLA 1230
Db 1009 LVLAIGRGVHTVHLCLGFVTSIGC---KRRASLALES-VLAPV---VHGALAAALA 1059
Qy 1231 KAQL-----IQIFFRNLITLGLLHGLVFLPVILSYGCP---DVNPALALEONRAE 1281
Db 1060 ASMLAASECGFVARLFLRLDLDIVFLGLIDGLLFFPILVLSILGPAAEVRPI----- 1110
Qy 1282 EAVAAMVASCPNHRSVST 1301
Db 1111 -----EHPERLST 1118

RESULT 56

US-09-754-032-4

; Sequence 4, Application US/09754032

; Publication No. US20030148368A1

; GENERAL INFORMATION:

; APPLICANT: SCOTT, MATHEW P

; GOODRICH, LISA V

; JOHNSON, RONALD L

; TITLE OF INVENTION: Patched Genes and their Use

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSER: Flehr, Hobbach, Test, Albritton & Herbert

; STREET: Four Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: US

; ZIP: 94111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/754,032
; FILING DATE: 03-Jan-2001
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1311 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-754-032-4

Query Match 7.6%; Score 526.5; DB 3; Length 1311;
Best Local Similarity 22.2%; Pred. No. 9.3e-36;
Matches 258; Conservative 175; Mismatches 420; Indels 307; Gaps 46;
Qy 336 LLGQFPQGGTGWASWPLTILVILSVIPVVALAAGLVFTLTDPVFLWNSAPNSQARSEKA 395
Db 72 ILGCLFQSDAGKLP--VAIILSTFCV-----GLKSAQIHRVDQLWVQEGRLAEALK 124
Qy 396 FHDQHPGP-FTNOVILTAPNRSSRYDLSLLGPKNFGSILDLDDLLELLELLELLELRLHL 454
Db 125 YTAQALGEADSTHQLVI---QTAKDPVSLLHP--GAL-----LEHL 162
Qy 455 QWMSPEAQRNI-----SLQDICYAPLNPD-----NTSLYD-----CCINSILQYFQNN 497
Db 163 KVVHAATRVTVHMYDIEMRLKDCYSPIDPEGYHHIESIIDNVIPCAIITPLDCFWEG 222
Qy 498 RTLL-----LTANOTLMGQTSQ--- 515
Db 223 SKLLGPDYPIYVPHLKHKLQWHLNPLVYVEEVKKLKQFPPLSTIEAYMKRAGITSAYMK 282
Qy 516 ---VDWKDHFLYC-ANAPLTFKDGATLALSCWADYGAPVFPF-----LAIGG----- 558
Db 283 KPCLDPTD--HCPATAP-NKSGHI PDVAELSHGCGYGAFAAAYMHMPQLIVGGATRNS 339
Qy 559 ---YKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLMEEAFLBENRPAQRMAQMFQVT 614
Db 340 TSALEKAKLQTVVQLMGEREMVEYADHYKHQVQKQKAAAVLDQAKRFAAEVRKI 399
Qy 615 FT-----AERSLEDEINRTAEDLPFATSYIVIFLYISLALGYSWSRVM 661
Db 400 TTSGSVSSAYSFPYFSTSTLNDILGKSEVSLKNIILGYMFLIYAVTL---IQW-RDP 455
Qy 662 VDSKATILGGLGAVVILGAVMAAGFFSYLGIRSLVILQVPPVLVLSVGADNIPFVLEY 721
Db 456 IRSQAGVGIAGVLLSITVAAGLGFALLGIPFNASQTQVPPFLALGLGVQDMFLTHY 515
Qy 722 QRLPRRGEPREVHIGALGRVAPSMCLLSLSEAICFFLGCALTPMPAVRTPALTSGLAVI 781
Db 516 --VEQAGDVPREERTGLVKKSGSLVSLASLCNMVAFALAAALLPAPAFVFCQAAILL 573
Qy 782 LDFLLQMSAFVALLSLDSKQEARSLDVCCVQPOELPP----- 820
Db 574 FNLGSILLVFPAMISLDRRSARADLLCLLPSPPLPKKIPERAKTRKNDKTHRIDT 633
Qy 821 -----PGQEGEGL-----LGFQKAYAPPLHMTIRGVVLLFLALFGVSLYSMC 865

QY 962 -TPSCRLYISGNKDKCPSTVNSLNCNCKMCSITMGSVRPSVEQ----- 1007
DB 787 EVASGCI-----TQEYWCNKASDE-GILAYKLMVQTGHVDNPIDKSLITAGHRLVDKD 838
QY 1008 -----FHKYL-PWFLNDRPNKCPKGGI-----AAYSTSVNLTSDGOVLASRFM 1050
DB 839 GIINPKAFNYLSAWATDALAGASQGNLKPQORWIHSPEVHLEIKSSPLIYQLP 898
QY 1051 AYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFPPYTTITNVFYEQYLITLP 1110
DB 899 FYLSGLSDTXSIKTLRSVRDLCLKYEA--KGLPN-----PFSGIPFLFWEQYLYLRT 949
QY 1111 EGLFMLSLCLVPFVASCLLGLDLSGLNLISIVMILVDTVGFMAWDISVNAVSLIN 1170
DB 950 SLLALACALAAVF--TAVMVLLNNAWAVALVTALATLQLGLVNMALLGVKLSAMPAYL 1008
QY 1171 LVSAGVMSVEFVSHITRSFAISPKPTWLERAKATISMGSAVPAVAMTNLRGILVLGLA 1230
DB 1009 LVLAIGRGVHTVHLCGLFVTSIGC-----KRRASIALES-VLAPV-----VHGALAAALA 1059
QY 1231 KAQL-----IQIFFRNLNLTLLGLLHGLVFLPVILSYVGP--DVNPALALEOKRAE 1281
DB 1060 ASMLAASECGFVARLFLRLLDIVFLGLIDGLLFPFIVLSILGPAAEVRPI----- 1110
QY 1282 EAVAAMVAVASCPNHPRSVST 1301
DB 1111 -----EHPERLST 1118

RESULT 58
US-10-369-493-5464
; Sequence 5464, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5464
; LENGTH: 1405
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5464

Query Match 7.4%; Score 514; DB 4; Length 1405;
Best Local Similarity 20.2%; Pred. No. 1.3e-34;

Matches 267; Conservative 190; Mismatches 412; Indels 452; Gaps 52;

QY 321 KGTSLSDKLSFSTHTLLGQFFQGWTVV--ASWPLTILVLSVIVVVALAAGLVFTLTD 378
DB 100 RKATGNRYALYSRLIQKLLFALGNVTHRNAWSI-ILAVSMIFAVC-CYGLQYVHIETD 157
QY 379 PYELMSA-----PNSQARSEKAFHQHGFPPF-RTN-----QVILTAPNRS 419
DB 158 IVKLWAQGRDLDEELNPLNPKAERNVTGDS--GPFLPRENGLGGVQVLIQTPYEBG 215
QY 420 YRYSLLGPKNPSGILDLDLLELLELQERLHRHQVSPQARNI-----SLQDICY 472
DB 216 --QDALAAGP-----LLKHVEIMKHIASFNVSVHGVDWSLSDICF 253
QY 473 APLNPD-----NTSLYD-----CC-INSLLOVFQNNRTL-----LIL 503
DB 254 KPAPPSVAADSAASSLGVDIVKIVPCIWITPIDCFWEGSKALGPHPSLPKSSLGPLGMILL 313

QY 504 TANQTMGQTSQVMDKD-----HFLYCANAPLTF----- 532
DB 314 SS-----LSDGDMIRMSDDPPIAVIDEIHRSFNLGSHYTFPERAGVSHGMDRPCIDPLDP 369
QY 533 -----KGTALALSCMADYCAPVPPFLAIGYKGYKDYSEAE 568
DB 370 ECPPMANVFDVCPHIDRVREIAKGYTELEBEKKKSGYSPPDFL--GRKKREAGDQ 426
QY 569 ALI----- 571
DB 427 KMHPAQPADSIPTIEDAVPAQVPVUSTAPIPTITTLSPPEARAAEKEKKQKARELKDYC 486
QY 572 -----MTVSLNNYP-----AGDP 584
DB 487 KSYRSAPFELWKKKDKWPEVMSENMYQNVDDYAAEMTGCCSGFASNVLNWPMEDMILGNP 546
QY 585 RLAAQ--AKL-----WEEATLEB-MRA 602
DB 547 RRKKGKGLSGADALQSFLVASPADVFLRFQKPGGRNSMKTGLDMDAMNETAAEQVLQA 606
QY 603 FORMMAGMPQVTFTAERSL-----EDEINRTTAEDLPFATS-----YIV 642
DB 607 WQR-----NFT--KSLYNHKANVEDGNERRTLH--PLASTSIADMLEEFCQNYTI 654
QY 643 IF-----LYISLALGSYSSWSRVWDSKATLG--LGGVAVVLGAVMAAMGFFSYLGRSSL 696
DB 655 ILAGYALMAYAVITQARFDNCLPATESSMGLALAGLVVTPASVAGLATWFGIEFNA 714
QY 697 VILQVVPFLVLSVGADNIFIVLEVQRLPRRRCEPREVHIGALGRHVARSMLLCSSEAI 756
DB 715 ATTQIVPFLTLTGIGVDMFMFLHNYRDVVYKLAGGHAEMAI--LMRETGMSILCTSNITL 772
QY 757 CFPLGALTMPMPAVRTFALTSLGLAVLTDLFLLOMSAFVALLSLDSKROEASRLDVCCCV--- 813
DB 773 SLTGTLLPPIALRSFCAQSSILLTFNFIAITITPAILISIDLRKKAORRDLVCLLYGD 832
QY 814 -----XPQBLPPPQGGEGL-----LLGFFQKAYA 837
DB 833 TREESYSMISKPKIOSKRIIGAPSEASIMQQFDGITQAQMASSDDDPAPWSLSHFIRYVI 892
QY 838 PELLHWITRGVVLLFLALFGVLSYMSCHISVGLDOELALPKDSYLLDYLFLNRYFEVG 897
DB 893 PFISKASKVAIIVGCCALLGASFIMGRQSTLGLGLDVLPEHTAPAQFLRADKTFSP- 951
QY 898 APVYVFTLTGYNFS--SEAGMNAICSSAGCNFESFTOKIQVATEPEQSILAIAPASSWV 954
DB 952 YPM-FAVINGPNIDYAHQORQIDNTRQSIGSSKYVKNK--NEEPSEKY-----WL 999
QY 955 DDFIDMLTPSSCCRLYISGNPKDKF--CPSTV-----NSLNC-----LKNCMSI 996
DB 1000 GLMRDWLI--SIQRGFDBEVAKGSFNLTSGTVIGSNVSEDAHLAHLAMCSHSLFCAG- 1056
QY 997 TMGSVR-----PSVEQPHKYL-PWFLNDR-----PNIKCPKGLAAYSTS 1035
DB 1057 RVGKIRLVDAAGIINSDFNYLTAFNVDHMYTVSQASFPPTP-----PKWELSKNHT 1112
QY 1036 VNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVRPY 1095
DB 1113 NEIPAAEPLAYSQIPFYLTGLTDTTAVIDAKIRSVCEP-----TQOGLPNFPQ 1163
QY 1096 TITNVFYEQYLITLPEGLFMLSLCLVPFVASCLLGLDLSGLNLISIVMILVDTVGF 1155
DB 1164 GIAFTFWEQYPLTGNLMAQISIIITISVFCVLSVLL-FNPMAALMVVCIILGIMTCELAF 1222
QY 1156 MALWDISVNAVSLINLVAGVMSVFEVSHITRSPAISTKPTWLERAKENATISMGSAVPAG 1215
DB 1223 MGLVGIKLPNSAVTLITAVGIGVEFTVHVVSFL-----TALGTRSORTSSAVDRVFPV 1277
QY 1216 V---AMTNLRGTLVLGLAKAQLIQIFFERLNLITLGLLHGLVFLPVILSYVGP--DVN 1270
DB 1278 VHGFSFTLLGILMLGSEFEPVVKYFFIVMTALICIGILINGLILLPVLLSVLSSRRRRTS 1337
QY 1271 P 1271

Db 1338 P 1338

RESULT 59

US-09-898-533-3

; Sequence 3, Application US/09898533

; Patent No. US20020106656A1

; GENERAL INFORMATION:

; APPLICANT: Gemmill, Robert M.

; APPLICANT: Drabkin, Harry A.

; TITLE OF INVENTION: TBC8, A GENE RELATED TO THE HEDGEHOG RECEPTOR,

; TITLE OF INVENTION: PATCHED

; FILE REFERENCE: 93445-00004

; CURRENT APPLICATION NUMBER: US/09/898,533

; CURRENT FILING DATE: 2001-07-02

; PRIOR APPLICATION NUMBER: US/09/268,140

; PRIOR FILING DATE: 2000-03-12

; NUMBER OF SEQ ID NOS: 46

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 3

; LENGTH: 1286

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

US-09-898-533-3

Query Match 6.9%; Score 474; DB 3; Length 1286;

Best Local Similarity 20.6%; Pred. No. 3.4e-31;

Matches 246; Conservative 176; Mismatches 465; Indels 306; Gaps 43;

QY 321 KGTSLDKLSFTHLLGQFGQGTWASWPLTILVLSVIPVVALAGLVTELTDPV 380

Db 43 KGKGRSRTAIYLSRVFSQSHLETGSSVQKHAGKVLFAILLVSTFCVGLKSAQIHSKVH 102

QY 381 ELWAPNSQARKEKAFHDQHG-PPERTNQVILTA---PNRSRYRDSLLGLPKNFSGIL 436

Db 103 QLWIEGGRLAEALAYTKTIGEDSATHQLLIQTHDNASVLHPQALL-----A 153

QY 437 DLDLLELELEQLERHLQVSPQAEORNTSLQDICVAPLNPDNTSLY-----DCC 486

Db 154 HLEVLKATAVKVHLVDTE-W-----GLRDMCMNPSTPGFEGIYIEQILRHLIPCS 204

QY 487 INSLOYFQN-----NRTLLLTAN-----QTLGQTSQVMDKH 521

Db 205 IITPLDPCFEGSGLGPEASVIVPGLNQLRLLMTLNPASVMQMKMSEKISDFETV 264

QY 522 FLY-----C-----ANAPLTFKD-----GTALALSQWADYGPVFPF 553

Db 265 EQMKRAAIGSGYMEKPCNLPLNCPDTPAKNKTQPPDVGNALLSGGCVG-YAAKHMW 323

QY 554 ---LAIGGYK---GKDYSAEALIMTFSL-----NNYPAGDPRLAQAK----- 590

Db 324 PEELIVGGRKRNRSGLHRAQALQSVVQLMTEKEMDQMDNKKVHHLGWTEKAAAEVLN 383

QY 591 LWEEAP---LEEMRAFORWAGHGFQVTFABRSLEDEINRTTAEDLPFATSIVIVFLYI 647

Db 384 AWQNFSSREVEQLLRKQSRATYDIYVFSAAALDDILAKFSPHSALSIVIGVAVTVLY- 442

QY 648 SLALGSYSSMSRYMVDKATLGLGGVAVLGAVMAGFFSYLIGRSSILVILQVFPFLVL 707

Db 443 --AFCTLLRW-RDPVQGSSVGAVGLLMCFSTAAGLGLSALLGIYFNAASTQVFPFLAL 499

QY 708 SVGADNIFIVLEBYQLRPRPGPREVHIGRALGRVAPSMMLCSLSEAICFFIGALATPMP 767

Db 500 GLGVGDHIFMLTAAYAESNR--EOTKL-----ILKVGPSILFSACSTAGSFFAAATPVP 553

QY 768 AVTFTALTGLAVILDFLLQMSAFVALLSDSKQEAERLDV-CCCV-----KPEQLPP- 820

Db 554 ALKVFCLOAAIVMCSNLAAALVFPAMISLDLRRRTAGRADIFCCCFVWMKEQPKVAPPV 613

QY 821 -----PGQEGEL-----LGFFQKAYAPP 839

Db 614 LPLNNNGRGARHPKSCNNRVRPLPAQNPLEQADIPGSSHSLSASFLATFAFQHTFPF 673

QY 840 LLHWITRGVLLFLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLNRYFEVGAP 899

Db 674 LMRSWVFLTWGFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTRLF--GFY 731

QY 900 VYFVTTLYNFSSEAGMNAICSSAGCNFSTQKIQYATEPEPQSYLAIP----- 949

Db 732 SMYAVTQ3-----NFEYPTQQLLRDY-HDSFVRVPHVKNKNGGL 771

QY 950 ASSWDDFIDWLTSSCCRLYISGPNKDKPCFSTVNSLCLKNC-----MS 995

Db 772 PDFWLLLFSEWL-----GNLQKIFDEYRDGRLTKECWFPNASSDAILAYKLI 819

QY 996 ITMGSVRPSVEQ-----FHKYL-PWFLNDRDNKPCPKGGLAAYSTS 1035

Db 820 VQTHVDNPDVKELVLTNRLVNSGDIINQRAFNYVLSAWTND-----VFAYG-- 867

QY 1036 VNLSDGQVLAASRFMAYHKLNKNSQD-----YTEALRAARELANITAD 1079

Db 868 ---ASQKLYPEPRQYFHP--NEYDLKIPKSLPLVYAQMPFYLHGLTDTSQI-KTLIGH 921

QY 1080 LRKVPGTDPAFEV--FPYITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLGLDLS 1137

Db 922 IRDLSVKYEGLFPNYPGIPFIPEQYMT-LRSSLAMILACVLLAALVLSLLLSVWA 980

QY 1138 GLMLLSITVMILVDTVGPMALWDISYNAVSLINLIVSVMGVSVEFVSHITRSPFISTKPTM 1197

Db 981 AVLTVLSVLASLAQIFGANTLLGKLSIPAIVALLISVGMMLCF--NVLSLGFMTSVGN 1038

QY 1198 LERAEATISGSAVFAGVAMTNLPGIILVGLAKAQLIQIPFRNLNLTLLGLLHGLVP 1257

Db 1039 RQRVQLSMQKSLGLVHGLMTSGVAVFMLSTSPPEFVIRHPCWLLLVLCVGCACNSLIV 1098

QY 1258 LPVILSYVGPVNPALALEOKRAEAAVAVMVASCPNHPSPRVSTADNIVNHS 1310

Db 1099 FPLLMSVGPAAE-LVPLE-----HPDRISTPSPUPVRSS 1132

RESULT 60

US-11-097-143-4068

; Sequence 4068, Application US/11097143

; Publication No. US20050208558A1

; GENERAL INFORMATION:

; APPLICANT: Venter, J. Craig

; APPLICANT: et al.

; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID

; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE

; TITLE OF INVENTION: DROSOPHILA GENES.

; FILE REFERENCE: CL000728

; CURRENT APPLICATION NUMBER: US/11/097,143

; CURRENT FILING DATE: 2005-04-04

; PRIOR APPLICATION NUMBER: 60/157,832

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: 60/160,191

; PRIOR FILING DATE: 1999-10-19

; PRIOR APPLICATION NUMBER: 60/161,932

; PRIOR FILING DATE: 1999-10-28

; PRIOR APPLICATION NUMBER: 60/164,769

; PRIOR FILING DATE: 1999-11-12

; PRIOR APPLICATION NUMBER: 60/173,383

; PRIOR FILING DATE: 1999-12-28

; PRIOR APPLICATION NUMBER: 60/175,693

; PRIOR FILING DATE: 2000-01-12

; PRIOR APPLICATION NUMBER: 60/184,831

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: 60/191,637

; PRIOR FILING DATE: 2000-03-23

; NUMBER OF SEQ ID NOS: 43008

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4068

; LENGTH: 1286

; TYPE: PRT

; ORGANISM: DROSOPHILA

Db 43 KGKAGSRRTAIVLRVFSQSHLETLGSSVQKHAGKVLFAVILVLSVFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHDQHFQ--PPFRTNQVILTA---ENRSSYRYDSLLGPKNFSGIL 436
Db 103 QLWIOBGGGLEAEALAVTQKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY 437 DLDLLELLELQERLRLHQLVWPEAQORNSLODI CYAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVLHYDTE-W-----GLRDMCNMPSPTSPFEGIYYIEQILRHLPIS 204
QY 487 INSLLOYFQN-----NRTLILLITAN-----QTLMGQTSQVDWDKH 521
Db 205 IITPLDCFWEGSOLLGPESAVVPIGLNQLRLTTLNPASVMQYKQKMSSEKISDFETV 264
QY 522 FYCANAPITFKDGTALALSCM-----ADYGAPVFPFLAIG--GYK 560
Db 265 EQYMKRAAIA---SGYMKPCLNPLNCPDPTAPNKNSTQPPDVG-----ILSGGCGYA 317
QY 561 GKDYSEAEALIMTFLSNYPAGDPLAQA-----589
Db 318 AKHMHWPPEELIVGGAKRN--RSGLRKAQALQSVMQVLTMEKEMYDQWQNYKVHHLGWTOE 376
QY 590 -----KLWEEAF---LEEMRAFORRMAGMFOVTFTAERSLEDEINRTTAEDLPIPATSY 640
Db 377 KAAEVUNAWQRFNSREVEQLLRKQSRITATNYDIYFSSAALDDIILAKFSPSALSIVIGV 436
QY 641 IVIFLYISALGSYSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFYSYLGIRSSILVILQ 700
Db 437 AVTVLY---AFCTLLRW-RDPVRGQSSVGVAGVLLMCFSTAGLGLSALLGIVFNAASTQ 492
QY 701 VVPLVLSVGAADNIFIVLEVQRPLRPGEPREHVHIGRALGRVARSMLLCSISEAICPPL 760
Db 493 VVPLFALGLGVHDHIFMLTAAYAESNR--EQTKL-----ILKVGPSILFSACSTAGSFFA 546
QY 761 GALTMPAVRTALTSLGLAVILDFLQMSAFALLSLDSKQESRLDV-CCCV-----K 814
Db 547 AFAPVPALKVFCLOAAIWMCSNLAAALLVFPAMISLDLRRRTAGRADIFCCCFPVWKEQ 606
QY 815 PQELPP-----PQOGBGL-----LLGFF 832
Db 607 PKVAPVPLNNNGRGARHPKSCNNRVALPAQNPLLEQRADIFGSSHSLSASFSLATFA 666
QY 833 QKAYAPFLHWTGRGVVLLFALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLFLNR 892
Db 667 PQHYTPFLMRSVWKFLTMVGMFLAALISLYASTRLQDGLDIIDLVPKDSNEHKFLDAQTR 726
QY 893 YFEVGAPVYVTLGVNFSSEAGMAICSSAGCNPFSTQKIQAYATEFPEQSYLAIP---- 949
Db 727 LF--GFYSMYAVTQG-----NFEYPTQOQLRDY-HDSFVRVPHVI 764
QY 950 -----ASSWDDFDIDLTPSSCCRLYISGNPKDKFCPTVNSLNLKNC----- 993
Db 765 KNDNGGLPDPFMLLSEML-----GNLQKIFDEERYDGLTKECWPPNASSDA 812
QY 994 -----MSITMGSVRPSVEQ-----FHKYL-PWFLNDRPNIKCPKGG 1028
Db 813 ILAYKLIVQTGHVDNPDVKELVLNRLVNSDGLINQRAFYNVLSAW-----ATNASSPTL 868
QY 1029 LAA--YSTSVNLTSQDVLASRFMAYHKPLNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLYPEPRQYFHPQ--NEYDLKPKSLPLVYVQAQMPFYLHGLTDS 926
QY 1071 ELAANITADLRKVPGTDAFEV--PEYTTITNVEQYLITLPEGLFMLSCLVPTPAVSC 1128
Db 927 QI-KTLIGHIRDLVKYEGFGPNYPSGIPFIFWEQYMT-LRSSGLAMILACVLLAALVLV 984
QY 1129 LLGLDLRSGLLMLLSVILMVLVDVTFGMALWDISYNAVSLINLVSAGVMSVBFVSHITRS 1188
Db 985 SILLLSVWRAVLVLSVSLASLAQIFCAWTLGLKLSAIPAVILILSVGMMLCF--NVLS 1042
QY 1189 FAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQIQTFFFLMLLITL 1248
Db 1043 LGFMTSVGNRRVQLSMQSLGPLVHGMLTSGVAVFMLSSTSPFEPVIRHFCWLLILVLC 1102

QY 1249 LGLLHGLVFLPVLSYVGPDPNPALALEQKRAEAAVAVMVASCPNHPRSVSTADNIYVN 1308
Db 1103 VGACNLSLVFPILLSMVGPEAE-LVPLE-----HPDRISTPSPLVR 1143
QY 1309 HS 1310
Db 1144 SS 1145

RESULT 63

US-10-835-517-48
; Sequence 48, Application US/10835517
; Publication No. US20050004028A1
; GENERAL INFORMATION:
; APPLICANT: Ingham, Phillip W.
; McMahon, Andrew P.
; Tabin, Clifford J.
; TITLE OF INVENTION: Vertebrate Embryonic Pattern-Inducing
; Proteins and Uses Related Thereto
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/835,517
; FILING DATE: 28-Apr-2004
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,771
; FILING DATE: 20-OCT-1997
; APPLICATION NUMBER: US 08/462,386
; FILING DATE: 05-JUN-1995
; APPLICATION NUMBER: US 08/435,093
; FILING DATE: 04-MAY-1995
; APPLICATION NUMBER: US 08/356,060
; FILING DATE: 14-DEC-1994
; APPLICATION NUMBER: US 08/176,427
; FILING DATE: 30-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: HMV-006.11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1299 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-10-835-517-48

Query Match 6.7%; Score 464.5; DB 5; Length 1299;
Best Local Similarity 20.4%; Pred. No. 2.3e-30;
Matches 245; Conservative 174; Mismatches 472; Indels 311; Gaps 41;
QY 321 KGTSLSDDLSPSTHTLLGQFFQCGWGTWVASWPLTILVLSVIPVWALLAAGLVFTELTPDV 380
Db 43 KGKAGSRRTAIVLRVFSQSHLETLGSSVQKHAGKVLFAVILVLSVFCVGLKSAQIHSKVH 102
QY 381 ELWASPNQARSEKAFHDQHFQ--PPFRTNQVILTA---ENRSSYRYDSLLGPKNFSGIL 436

Db 103 QLWIOGGGLEAEAYTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
Qy 437 DIDLLELLELQERLRLHQLVWSPQAQRNLSLDICVAPLNPNTSLY-----DCC 486
Db 154 HLEVLVKATAVKVLHYDTE-W-----GLRDMCMPSPTSPSGIYIEQILRLHLPSCS 204
Qy 487 INSLLOYFON-----NRTLALLLTAN-----QTLMGQTSQVDWKDH 521
Db 205 IITPLDCFEGSQGLQGPESAVVIPGLNQRLWTLTLPASVMQVMQKQSEEEKISIDFETV 264
Qy 522 FLYCANAPLTFKDGXTALASC-----ADYGAPVPPFLAIG--GYK 560
Db 265 EQVMKGAATA--SGYMERPCPLNPLNCPDTPAPKNKSTQPPDVG-----ILSGGCGYA 317
Qy 561 GKDYSEBALIMTFSLNYPAGDPRLAQA----- 589
Db 318 AKHMHVPEELIVGGAKRN--RSGHLRKAQALQSVMQTEKEMYDQMDNYKVHHLGWTOE 376
Qy 590 -----KLWEEAF--LEEMRAFORERAGMFWQVTFTAERSLEDEINRTTAEDLPIPATSY 640
Db 377 KAAEVUNAWQRFSPREVEQLLRKQSRITATNYDIYVFSAALODILAKFHPALSIVIGV 436
Qy 641 IIVFLYISALGSYSSVMVDSKATLGLGGVAVVLGAVMAAMGFYSVLGRSSIVLIQ 700
Db 437 AVTVLY--AFCTLLRW-RDPVRGQSSVGAGVLLMCFSTAGLGLSALLGIVENASTQ 492
Qy 701 VPFVLVSGADNIFTVLEYQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEACFPPL 760
Db 493 VVPFLALGLGVDFHIFMLTAAYAESNR--EQTKL-----ILKKVGPSILFSACSTAGSFFA 546
Qy 761 GALTMPAVRTFALTSLGAVILDFLQMSAFVALLSLDKROEASRLDV-CCCV-----K 814
Db 547 AAFIPVPAKVFCLQAAIWMCSMLAALIVFPAMISLDLRRRTAGRADIFCCCFPPWKQ 606
Qy 815 PQELPP-----PQGEGL-----LLGFF 832
Db 607 PKVAPPVPLNNNGRGARHPKSCNNRVALPAQNPLLEQRADIPGSSHSLASFSLATFA 666
Qy 833 QKAYAPFLHWITRGVLLFLALFGVLSYKCHISVGLDQELALPKDSYLLDYFLFLNR 892
Db 667 FQHYTFPLRWSWKFLTVNGFLAALISLYASTRLQDGLDIDLVPKDSNEHKFLDAQTR 726
Qy 893 YPEVGAPVFTVTLGYNFSSEAGMNAICSSAGCNFESFTQKIQYATPEPQSVLAP--- 949
Db 727 LP--GFSYKAVTQG-----NFEYPTQQLRDY-HDSFVRVPHVI 764
Qy 950 -----ASSWVDDFDLWLTSPSCCRLYISGNKOKFCFSTVNSLCLNKC----- 993
Db 765 KNDNGGLPDFWLLLFSEWL-----GNLQKIFDEEYRDGRLTKECWFPNASSDA 812
Qy 994 -----MSITMGSVRPSVEQ-----PHKYL-PHFLNDRPNIKCPKG 1028
Db 813 ILAYKLIQVQGHVDNPDVKELVLTNRNVNSDGIINQRAFNYLSAW-----ATNASSPTL 868
Qy 1029 LAA--YSTSNLTSQVLSRPMAYHKPLKNSQD-----YTEALRAAR 1070
Db 869 LRANCIRNRANGASQGLKPEPRQYFHPQ--NEYDULKIPKSLPVVTAQMPFVLHGLTDS 926
Qy 1071 ELAANTADLRKVPDPAFEV--FPYTTITNVEQYLATLPEGLFMLSCLLVPPTFAVSC 1128
Db 927 QI-KTLIGHIRDLVSVKYEGFGLPNYPSPGIPFIWEQYMT-LRSSLAMILACVLLAALVIV 984
Qy 1129 LILGLDLRGLNLLSIVMLIVDTVGFMAWDNISYNASVLINLVSAGHVSFVSHITS 1188
Db 985 SLLLSVMAAVLVLSVLAASLAQIFGAMTLGLKLSAIPAVILILSVGMMLCF--NVLIS 1042
Qy 1189 FAISTKPTWLERAKEATISMSAVFAGVAMTNPGLVLGLAKAQIQLIFFFLNLILITL 1248
Db 1043 LGFWTSVGNRRQVQLSQMSQSLGPLVHGMLTSGVAVFMLTSPFPFVIRHFCWLLLVJVC 1102
Qy 1249 LGLHLGLVFLPVLTSYVGPVNPALALEQKRAEEAAVAVMASCNPHSPSRVSTADNIYN 1308
Db 1103 VGACNSLLVFPILLSMVGPEAE-LVPLE-----HPDRISTPSPLPVR 1143

Qy 1309 HS 1310
Db 1144 SS 1145

RESULT 64

US-10-890-776A-4809
; Sequence 4809, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Aecomica Sequence Listing Engine
; SEQ ID NO 4809
; LENGTH: 906
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
; FEATURE:
; OTHER INFORMATION: Sequence of the Patched motif of gi 3892144
US-10-890-776A-4809

Query Match 6.4%; Score 442.5; DB 5; Length 906;
Best Local Similarity 20.0%; Pred. No. 1.1e-28;
Matches 202; Conservative 196; Mismatches 367; Indels 243; Gaps 39;

Qy 378 DPVELWSPNSQARSEKAFHQHGFPTFTNQVILTAPNRSRYRSDLSLLGPKNFSGILD 437
Db 5 DPSYVFTPSDARWRREISVFNENW-PL--DENKFLPGKSFKAERFVNILIRAKDGGSIMR 61
Qy 438 LDLLELLEL-QERLRLHQLVWSPQAQRNLSLDICVAPLNPNTSLYD--CCINSLLQYF 494
Db 62 DNVLHEIELNQMWNNISIPDTDLKFNLTQYDCLLS-----YDMVCCGANEHIQ-- 110
Qy 495 QNNRTLLILLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGXTALASCADYGA PVPFPL 554
Db 111 -----MLLRND--VNQILDL---HF-----P---RGGTK-----DTPVY--- 137
Qy 555 AIGYKGKDYSEAEALIMTFSLNYPAGDPRLAQA-----AKLWEE---AFLEEMRAP 603
Db 138 -LGGIFGD-----VQFF--QNGTSLDAKLTQLFYFLKQDQKQVVEEYSSKFSYALETF 186
Qy 604 QRKM--AGMFQVTFTAERSLEDEINRTTAEDLPIFATSYVIFLYI---SLALGSYSSWS 658
Db 187 LNQVYSSDVITLSFAHYQSLDGLDENAKAFVFNFWVSFFVLAMVALVSSFTLKSSSATK 246
Qy 659 RVWDTSKATLGLGGVAVVLGAVMAAMGFYSVLGRSSIVLIQVVPPLV----- 707
Db 247 IDWISSKPLAAAGMFPSTVLSIISAFGLFILGVRYN-VINTIIPFLIIGEFQSFQKMEK 305
Qy 708 ----SVGADNIFVLEQRLPRRPGEPREVHIGRALGRVAPSMMLCSLSEALCFFLGLAL 763


```
Db 306 NQIPALGIDMFLWNAWQTSKSLVPE--RMSKTLSHAGVAVTITNVTDVNSFAIGCI 363
Qy 764 TTPAVTRTALTSLGLVLDLFLQMGAFVALLSL-----797
Db 364 TDLPGIQFFCIYACVSFAFYFQLATFFSGAMAIMGEVEREKHCLPFYRTTQLVDISKM 423
Qy 798 -----DSKRQE-----ASRLDVC--CVKQELPPPG--822
Db 424 NEBADSQLOQIKRSASPAPFNFLSSNSFSFSDSDSFSSKKTIPAEPAWKEQQSPNSSL 483
Qy 823 -----QSGELLGFFQKAYAPFLHWTIRGVVLLLELALFGLVSLYSNCHISVGLDQEL 875
Db 484 SKSKDREKORLVHFIKIGYGFILNSVRIPSGLIFFVYLAIAMYGCNFRGLNPGN 543
Qy 876 ALPKDSYLLDYFLFNRYFEVGAFFVFTTGLGNFSSEAGMNAICSSAGCMMNPFQTKIQ 935
Db 544 LVTDHYIAKYFSDIKHFWRIGLAQLHAVLNPNTLISERNEL-----LK 589
Qy 936 YATEPPEQSYLAIPASWDDDFIDMTTPSSCCRLYISGPNKDKFCPSTVNSLCLNCKMS 995
Db 590 VVSFAFNTQY--TLGREGTVFFLLLEYLN-----YLSLNAE-----V8 625
Qy 996 ITWGSVRPSVEQPHKYLPHFLNDRPNIKCPKGLAAYSTSVNLT--SDGQVLASRFMAYHK 1054
Db 626 DTERLWTKLNSWLKY-----TGSGTQWASNLKINKTDGSGFQAFRQIALK 671
Qy 1055 PLKNSQDYTEALRAARELAN-----ITADLRKVPGTDPAPFVFPYITNV-----1100
Db 672 NFVEPNHKAQAALLADIADHPFNVVVHVSFGNRKIL--NDFPSSHSCYAKNIPKLA 730
Qy 1101 --FYEQYLITLP---EGLFMLSCLVPTPAVSCLLGLDLRSLGLNLALSIVMLVDVTGF 1155
Db 731 PFPADQYLLIPATIQQNVISLLCM-----AVVSFLLVPSLPSGFVIFVSIVSINIGVGY 786
Qy 1156 MALWDSYNVAVSLINLVSAGVSVFVSHITSPALSTPTWLERAKEATISMGSNAFAG 1215
Db 787 MTLGWNLDAVSMISITMSIGFAVDLSAHIYAFVTSBGDT--KORVIGALETLGWPIFOG 845
Qy 1216 VAMTNLPGLVLGLAKAQIQPFRLNLILTLGLLHGLVFLPVILS 1263
Db 846 ASST--IAGSILYTDVAYII--LVFFKTIWLTWMLGAIHGLFPIPIPLS 891

RESULT 65
US-08-954-701A-6
; Sequence 6, Application US/08954701A
; Publication No. US20030032085A1
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley, Hoag & Eliot
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII(text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/954,701A
; FILING DATE: 20-OCT-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36709
```

```
; REFERENCE/DOCKET NUMBER: SUV-003.08
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1285 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-954-701A-6

Query Match 6.3%; Score 438.5; DB 2; Length 1285;
Best Local Similarity 20.0%; Pred. No. 4.3e-28;
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;

Qy 321 KCTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSPVVALAAGLVFTELTTDPV 380
Db 43 KKGARGSRTAIYLRVSFQSHLETGLSSVQKHAGKVLFAILVLSTFCVGLSKSAQIHSKVH 102
Qy 381 ELWSAPNSQARSEKAFHQHFG--PFPRTNOVILTA---ENRSSYRVDSLLGLPKNSGIL 436
Db 103 QLWIOEGGRLERLARLAYTKTIGEDSATHQLLIQTHDPNASVLHPQALL-----A 153
Qy 437 DLDLLELELEQLERLRLHQVNSPEAQRNLSLQDICYAPLNPDNTSLY-----DCC 486
Db 154 HLEVLKATAVKAVHLYDIE-W-----GLRDMCNMPTSPFEGIYYIEQILRHLPIS 204
Qy 487 INSLIQQYFON-----NRTLLLLLTAN-----QTLMGQTSQVDMKDH 521
Db 205 IITPLDCEWEGSOLLGPESAVVIPGLNQRLLTTLNPASVMQVMKQMSSEKISDFETV 264
Qy 522 FLY-----C-----ANAPLTFKD-----GTALALSQWADYGAPVFPF 553
Db 265 EYMKRAAIGSGYMKPCPLNPLNCPDTPAPKNSTQPPDVGAILSGCYG--YAAKMHMW 323
Qy 554 ---LAIGGYK---GKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK-----590
Db 324 PEELIVGGKRNRSGLHKAQALQSVQLMTEKEMVDQMDNYKVHHLGWTQEKAEVLN 383
Qy 591 LWEEAF---LEEMRAFQRMAGMFOVTTAERSLDEINRTTAEDLPFATSYVIFLYI 647
Db 384 AMQRFNREVEQLLRKQSRITATNYDIYVFSSAALDILAKFHSHPALSIVIGVAVTVLY- 442
Qy 648 SLALGSYSWSRVMVDSKATLGLGVAVVLGAVMAAMGPFPSYLG------692
Db 443 --AFCTLLRW--RDPVRGQSSVGAVGLLMCFSTAAGLGLSALLGIVFNALTAAYAESNRR 499
Qy 693 --RSSLVI---LQVVPFLVLSVGADNIFIFVLEYQLPRRPGEPREVHIGRALGRVAPSM 747
Db 500 EQTKLILKNASTQVVPFLALGLGVDFHIFT-----VGPSI 533
Qy 748 LLCSLSEACFPFLGALTMPMPAVRTTALTSGLAVIDLFLQMSAFVALLSLDSKROEASRL 807
Db 534 LFSACSTAGSPFAAAPIVPALKVFCLQAAIVMCSNLAALLVFPAMISLDRRLRTAGRA 593
Qy 808 DV-CCCV-----KQELPP-----PQ 823
Db 594 DIFCCCFPVWKEQPKVAPPVPLPLNNNGRGARHPKSCNNNRVPLPAQNPLLEQRADIPGS 653
Qy 824 GEGE-----LLGPFQKAYAPFLHWTIRGVVLLFLALFGVSLYSNCHISVGLDQELALPK 879
Db 654 SHSLASFSLATPAFOHYTFFLMRSWVKFLTWGFLAALISSLYASTRLQDGLDIDLVPK 713
Qy 880 DSYLDYFLFLNRYFEVGAFFVFTTGLGNFSSEAGMNAICSSAGCMMNPFQTKIQAYATE 939
Db 714 DSNHKKFLDAQTRLF--GFYSMAVTVQG-----NFEYPTQQQLLRD 752
Qy 940 PPEQSYLAIP-----ASSWDDFDLWLTTPSSCCRLYISGPNKDKFCPSTVNSLNC 989
Db 753 YHDS--FRVPHVKNNDNGGLPDLFWLLFSEWL-----GNLQKIFDEYRDGRL 798
```

QY	990	LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db	799	TKCWFPNASSDAILAYKLI VOTGHVDNPDVKELVLTNRLVNSDGIINQRAFYNYLSAWA 858
QY	1016	LNDRPNIKCPKGGAAAYSTSVNLSDQVLASRFMAYHKPLKNSQD-----1061
Db	859	TND-----VFAYG-----ASQGLKYPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901
QY	1062	--YTEALRAARELANITADLRKVPGTDPAFEV--FPYTTINVFYEQVLTILPGLPMLS 1117
Db	902	PFYLHGLTDTSQI--KTLLIGHIRDSLVKVEGFLGPNYPSGIPIFWEQYMT--LRSSLAMIL 959
QY	1118	LCLVPTFAVSCLLGLDLRLSGLLMILLSIVMLIVDTVGPMALWDISYNNAVSLINLVSAGM 1177
Db	960	ACVLAALVSVLSSWAAVVLVLSLASLAQIFGATMLLGIKLSAIPAVILILSVGM 1019
QY	1178	SVEFVSHITRFAISTKPTWLERAKEATISGMSAVFAGVAMTNLPGLILVLGLAKAQLIQI 1237
Db	1020	MLCP--NVLISLGFTSVGNRRVQLSQMSGLGFLVHGMLTSGVAVFMLSTSPFEFVIR 1077
QY	1238	PPFRLNLITLGLLHGLVFLPVILSVYVGPVNPALALEQKRAEAAVAAMVASCNPNPS 1297
Db	1078	HFCWLLVVLVCGACNSLLVFPILLSMVGPEAE--LVPLE-----HPD 1118
QY	1298	RVSTADNIYVNH5 1310
Db	1119	RISTSPPLVRSS 1131
RESULT 66		
US-09-754-032-6		
Sequence 6, Application US/09754032		
Publication No. US20030148388A1		
GENERAL INFORMATION:		
APPLICANT: SCOTT, MATHEW P		
GOODRICH, LISA V		
JOHNSON, RONALD L		
TITLE OF INVENTION: Patched Genes and their Use		
NUMBER OF SEQUENCES: 19		
CORRESPONDENCE ADDRESS:		
ADDRESSEE: Fiehr, Hohbach, Test, Albritton & Herbert		
STREET: Four Embarcadero Center, Suite 3400		
CITY: San Francisco		
STATE: CA		
COUNTRY: US		
ZIP: 94111		
COMPUTER READABLE FORM:		
MEDIUM TYPE: Floppy disk		
COMPUTER: IBM PC compatible		
OPERATING SYSTEM: PC-DOS/MS-DOS		
SOFTWARE: Patentin Release #1.0, Version #1.30		
CURRENT APPLICATION DATA:		
APPLICATION NUMBER: US/09/754,032		
FILING DATE: 03-Jan-2001		
CLASSIFICATION: 435		
PRIOR APPLICATION DATA:		
APPLICATION NUMBER: US/08/540,406		
FILING DATE: 06-OCT-1995		
ATTORNEY/AGENT INFORMATION:		
NAME: Rowland, Bertram I		
REGISTRATION NUMBER: 20015		
REFERENCE/DOCKET NUMBER: a60190-1		
TELECOMMUNICATION INFORMATION:		
TELEPHONE: 415-781-1989		
TELEFAX: 415-398-3249		
INFORMATION FOR SEQ ID NO: 6:		
SEQUENCE CHARACTERISTICS:		
LENGTH: 1285 amino acids		
TYPE: amino acid		
STRANDEDNESS: single		
TOPOLOGY: linear		
MOLECULE TYPE: protein		
SEQUENCE DESCRIPTION: SEQ ID NO: 6:		
US-09-754-032-6		
Query Match		
Best Local Similarity 20.0%; Pred. No. 4.3e-28;		
Matches 242; Conservative 174; Mismatches 450; Indels 347; Gaps 44;		
6.3%; Score 438.5; DB 3; Length 1285;		
QY	321	KGTSLSDKLSFSTHTLLGGFPQGWGTVASWPLTILVLSVIPVVALAAGLVFTLTDPV 380
Db	43	KGKARSRTAIYLRVSFQSHLETGSGVQKAGKVLFAVLVLSTFCVGLKSAQIHSKVH 102
QY	381	ELWSAPNSQARSKAFHQHFG-PFRTNQVILTA---PNRSSYRYSLLGPKNFGSIL 436
Db	103	QLMTQEGRLAEALAVTKTIGEDSATHQLLIQTTHDPNASVLHPQALL-----A 153
QY	437	DLDLLELLEQLRHLRHLQWSPEAQBNISLQDICYAPLNPNTSLV-----DCC 486
Db	154	HLEVLVRATAVKVHLYDTE-W-----GLRDMCNMPSPTSPFEGIYIIEQILRHLRILPCS 204
QY	487	INSLTQVFN-----NRTLLLTAN-----QTLMGQTSQVDMKH 521
Db	205	IITPLDCFWEGSQLLGPESAVVIPGLNQLRLLWTLNPNASVMQMKQKSEKISFDFTV 264
QY	522	FLY-----C-----ANAPLTFKD-----GTALALSCMADYAGPVPPF 553
Db	265	EQYKRAAIGSGYMEKPCLNPLNPNCPDTPAPKNSTQPPDVGAILSGCYG-YAAKGMHW 323
QY	554	---LAIGGYK---CKDYSEAEALIMTFSL-----NNYPAGDPRLAQAK----- 590
Db	324	PEELIVGGRKRNRSGLRKAQALQSVQVQLMTEKEMYQWQDNKYVHHLGWTQEKAAEVLN 383
QY	591	LWEEAF---LEEMRAFQRRMAGMFQVTFTAERSLEDEINRTTAEDLPFATSYIVIFYI 647
Db	384	AWQNFSEVEQLLRKQSRITATNDIYVSSAALDDILAKFSHPALSIVIGVAVTVLY- 442
QY	648	SLALGSYSWSRVMDSKATLGLGSAVVLGAVVAVLGAVMAMGFFSYLGI----- 692
Db	443	--APCTLLRW-RDPVRGQSGVGVAGVLLMCFSTAAGLGLSALLGVFNALTAAYAESNR 499
QY	693	--RSSLVI---LQVVPFLVLSVGADNIFI FVLEYQRLPRRPGEPREVHIGRALGVAPSM 747
Db	500	EQTLLKLNASTQVVPFLALGLGVNDHIFI-----VGPSI 533
QY	748	LCLSLSEALCPFLGALTMPAVRTPALTSGLAVILDLLQMSAFVALLSLDSKQREASRL 807
Db	534	LFSACSTAGSFFAAAFI PVPALKVFCLOAAI WMCNLAALILVFPAMISLDRRTAGRA 593
QY	808	DV-CCCV-----KQELPP-----PCQ 823
Db	594	DIFCCCFPVWKEQPKVAPPVPLPNNNGRGARHPKSCNNRVPLPAQNPLLSQRADIPS 853
QY	824	GEGL-----LLGFFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVGLDQELALPK 879
Db	654	SHSLASFSLATFAFOHYTFLMRSVKFLTVNGFLAALISSLYASTFLQGLDILDLVPK 713
QY	880	DSVLLDYFLFLNRIFYEVGAPVFTVTLTGYNPFSEAGMAICSSAGCANNPSTFKIQIYATE 939
Db	714	DSNEHKFLDAQTLRP--GFYSMVAVTQG-----NFEYPTQQQLLRD 752
QY	940	FPEQSYLAIP-----ASSWDDFDITDWTTPSSCCRLYISGPNKDECPSTVNSLNC 989
Db	753	YHDS--FRVPHVIKNDNGGLPDPFWLLLFSEWL-----GNLQKLFDEYDRDGL 798
QY	990	LKNC-----MSITMGSVRPSVEQ-----FHKYL-PWF 1015
Db	799	TKCWFPNASSDAILAYKLI VOTGHVDNPDVKELVLTNRLVNSDGIINQRAFYNYLSAWA 858
QY	1016	LNDRPNIKCPKGGAAAYSTSVNLSDQVLASRFMAYHKPLKNSQD-----1061
Db	859	TND-----VFAYG-----ASQGLKYPEPRQYFHP--NEYDLKIPKSLPLVYAQM 901
QY	1062	--YTEALRAARELANITADLRKVPGTDPAFEV--FPYTTINVFYEQVLTILPGLPMLS 1117
Db	902	PFYLHGLTDTSQI--KTLLIGHIRDSLVKVEGFLGPNYPSGIPIFWEQYMT--LRSSLAMIL 959

Db 1020 MLCP--NVLSLGFMTSVGNRRQVQLSMQMSLGLVHGMMLTSGVAVFMLSTSPFBEVIR 1077
QY 1238 FFRLNLITLLGLHGLVFLPVLISVVGPDVNPALALEQKRAEEAFAVAVMVASCPNHP 1297
Db 1078 HFCWLLLVLCVAGNSLLVFPILLSMVGPAAE-LVPLE-----HPD 1118
QY 1298 RVSTADNIYVNH 1310
Db 1119 RISTPSPLEVRSS 1131
RESULT 68
US-10-493-493-5410
; Sequence 5410, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; PRIOR FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5410
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5410
Query Match 5.5%; Score 379; DB 4; Length 714;
Best Local Similarity 23.0%; Pred. No. 2.6e-23; Mismatches 299; Indels 160; Gaps 25;
Matches 173; Conservative 119;
QY 545 DYGAPEPFLAIGYKGDYSEAEALIMTFSLN--NYP----AGDPRLAQ--AKLWEAEAF 596
Db 58 DYAEH-----TGCGS-----FASVNLNWPEDMILGNPRRAKGGKLSGADA 100
QY 597 LEEMRAFORRMAGMFOVTTAERSLEDEINRTTAEDLPATSYIVIFLY----- 646
Db 101 LQ----FVFLVAPADVFLRFKQSLSE-----TLETLSLYEEHCILHRRHPLLCRSNSV 151
QY 647 ----ISLALGYSYSSWRVMV-----DSKATLGLGVAVVLGAVMAAGPFSL 690
Db 152 NSTIPSSLLDMLSCWMLLLILLDTTAFRLQHONGFALGVLVVTFASVAGIGLATWF 211
QY 691 GIRSSLVILQVPLVLSVGADNIFVLEYQRLPRRPGPREVHIGRALGRVAPSMLLC 750
Db 212 GIFENATTOIVFLTLGIGVDNFMFLHNYRDVVKLAGHAEWAI--LMRETMGSILCT 269
QY 751 SLSEAICTFFLALTPMPAVRTFALTGLAVILDFLLQMSAFVALLSDSKRQASRLDVC 810
Db 270 SINNLISFLGTLLPILALRSFCAQSILLTFNFIAITTYPAIISIDLERKKAQRDFD 329
QY 811 CCVK-POELPPPGGEGI-----LLGFFQKAYAPAFLLHMTTRGVLLLLFL 854
Db 330 TREESTSEASIMQOFGDITQAOVMASSDDPAPWSLHSFIRYIYPTFKPASKVAIIIGCC 389
QY 855 ALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYVVTTLGVNFS--- 911
Db 390 ALLGASFVGMRSQTLGLGLGDVLEHTAPAQFLRAQKYESF-YPM-FAVIKGNIDYAH 447
QY 912 SEAGNNAICSSAGCNNSFTQKIYATYEPFQSYLAIPASSWDDFTDMLTPSCCLYI 971
Db 448 QQRQIDNRYQSIGSSKVIKNK-----NEEPSEK-----WLGIMRDLII--SIQGF 494
QY 972 SGPNNKDF-----CPSTVNSLNCNCSITMGSVRPSVEQPHK--YLPFWFLNDRPNI 1022

Db 495 EEVAKGSFNLTSGTVIGSNVSEDARLAHALMCHSGSLFECAGRIGKIRLIPFYL----- 549
QY 1023 KCPKGGIAYSTSVNLTSDQVLASFMAFHKPLKNSQDYTEALRAARELAANITADLRK 1082
Db 550 -----GLTDTAVIADAIDRSVCERF----- 571
QY 1083 VPCTDAFEVFPYTTITNVFYEQYLTLPGLFMLSCLVPTFAVSCLLGLDLRSGLNL 1142
Db 572 ----TDQGLPNFPQGIATFWEQYLFGLTNLMQAISITISVFCVISVLL--FNPWAALMV 627
QY 1143 LSIWMLVDTVGMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAK 1202
Db 628 CILGIMTCELAFGLVGLIKLNPVSAVTLITAVGIGVEFTVHVVSFL-----TALGTRS 682
QY 1203 EATISMGSAVPAGV---AMTNLPGILVGLA 1230
Db 683 QRTSSAVDRVFPVPIHSGFSSTLIGILMLGFS 713
RESULT 69
US-10-060-756A-6
; Sequence 6, Application US/10060756A
; Publication No. US200300467117A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/060,756A
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 6
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-756A-6
Query Match 5.1%; Score 351.5; DB 4; Length 767;
Best Local Similarity 20.5%; Pred. No. 7.3e-21; Mismatches 285; Indels 189; Gaps 31;
Matches 158; Conservative 137;
QY 334 HT-----LLGQFFQCGMTWVASMPLTILVLSVIPVVALAAGLVF--TELTDPFVLSWAP 386
Db 115 HTDCLGELLSRTFQMLGWQVGAHPWIFLLAPLMLTAALGTGLYLPKDEEDLEEHVTPV 174
QY 387 NSQARSEKAFHDQHFQFPFRTNQVILTAPNRSS-----YRDSLLGLGPKNFSGL 436
Db 175 GSPAKAERFVQGH-----FTTNDYSRFSASRRSTEANFVSLVSVSYSDSLDDPATFAEVS 230
QY 437 DLDDLELLELQELRHLQVWSEARNSLQDLC--YAPLNPDNTSLYDCCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSQYQVQVCARYAL-----CVPPNPILY 270
QY 493 YFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKQGTALALSMAVGA-PVF 551
Db 271 AWQVNTKILNL-----SSISFPAY-----NHGRHPLY 296

```
QY 552 PFLAIGGY-----KGDYSEAEALIMTFSLNNYPAGDPRL-AQAKLWEEAFLEEMRAF 603
Db 297 LTGFFGGYILGSLGQGLLRKAMRLLYLK---TEDPEYDVQSKQMLTHLLDQFTNI 353
QY 604 QRMA--GMFQVFTTA-ERSLEDEINRTTAEDLPFATSIVIVFLYISIALGSSYSSWSRV 660
Db 354 KNILAKKLEVVHFTSLRQLEFATSVTV--IPVHLAYILILFVNTSCFRPD-----406
QY 661 MVDKATLGLGGVAVVLGAVMAAGPFYSYLGRSSILVILQVVPFLVSGADNIFIFVLE 720
Db 407 CIRNKMCAAFVISAFVAVSGFGLLHIGV-PFVIIIVANSFPLILGVGDDMFIMISA 465
QY 721 YQRLPRPCEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSGLAV 780
Db 466 WHK--TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSVQCFTYGMTL 523
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCKVKQELP-----PPGOG 824
Db 524 LFCYFYNITCFGAFWALDKREV-----VCLWLKKAADPKWSPKFKCCFPFGSVPDHGH 578
QY 825 EGL--LLGFPQKAYAPFLHWTGRVVLFLFALFGVSLYSCHISVGLDQELALPKDSY 882
Db 579 TDHPISLFRDYFGFPLTRSESKYFVVIYLYIISIIYGGCFHQBEGDLRLNLASDDSY 638
QY 883 LLDYFLFLNRYF-EVGAPVYFVTTLGYNFSSEAGNMAICSSAGCNFSTQKIYATEPP 941
Db 639 ITPYFNVEENYFSDYGRVMVITKVDYWDK-----DVRQKLENTKIP 683
QY 942 EQSYLAIP--ASSWVDDFDMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMG 999
Db 684 EKNVYVDKLTPEFLDAVQ-----YLGNSQD---PNEKNT-----717
QY 1000 SVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVNLTSDGVLASR 1048
Db 718 -----FMNIPDLSNFPN-----FQHDINISSNEIISR 748

RESULT 70
US-10-890-776A-6
; Sequence 6, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR APPLICATION NUMBER: US 10/060,756
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Ascomica Sequence Listing Engine
; SEQ ID NO 6
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-890-776A-6
```

```
Query Match 5.1%; Score 351.5; DB 5; Length 767;
Best Local Similarity 20.5%; Pred. No. 7.3e-21;
Matches 158; Conservative 137; Mismatches 285; Indels 189; Gaps 31;

QY 334 HT-----LLGQFQCGHTVWASWPLTILVLSIVPVVALAAGLVF--TELTDTDVLSWAP 386
Db 115 HTDCEGLLSRTFQWLGWQVGAHPWIFLLAPLMLTAALGTGLYLPKDEEDLEEHYTPV 174
QY 387 NSQARSEKAFHQHGFPRPTNOVILTAENRSS-----YRDSLLLGPKNFGSGL 436
Db 175 GSPAKARRFVQGH-----FTTNDSTRFSASRSTEAANFVSLVSVSDSLDPAFAVS 230
QY 437 DLDLLELELELERLHLQVWSPQAQRNISLQDIC--YAPLNPDMNTSLYDCCI--NSLLQ 492
Db 231 KLDGAVQDLRV-----AREKGSQIQYQVCARYEAL-----CVPPNPILY 270
QY 493 YFQNRRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKGTALALSCMADYGA-PVF 551
Db 271 AMQVNTKLTNL-----SSISFPAY-----NHGRHPLY 296
QY 552 PFLAIGGY-----KGDYSEAEALIMTFSLNNYPAGDPRL-AQAKLWEEAFLEEMRAF 603
Db 297 LTGFFGGYILGSLGQGLLRKAMRLLYLK---TEDPEYDVQSKQMLTHLLDQFTNI 353
QY 604 QRMA--GMFQVFTTA-ERSLEDEINRTTAEDLPFATSIVIVFLYISIALGSSYSSWSRV 660
Db 354 KNILAKKLEVVHFTSLRQLEFATSVTV--IPVHLAYILILFVNTSCFRPD-----406
QY 661 MVDKATLGLGGVAVVLGAVMAAGPFYSYLGRSSILVILQVVPFLVSGADNIFIFVLE 720
Db 407 CIRNKMCAAFVISAFVAVSGFGLLHIGV-PFVIIIVANSFPLILGVGDDMFIMISA 465
QY 721 YQRLPRPCEPREVHIGRALGRVAPSMILCSLSEALCFPLGALTMPAVRTTALTSGLAV 780
Db 466 WHK--TNLADDIRE-RMSNVYSKAAVSITITITNIALYTGIMSSFRSVQCFTYGMTL 523
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCKVKQELP-----PPGOG 824
Db 524 LFCYFYNITCFGAFWALDKREV-----VCLWLKKAADPKWSPKFKCCFPFGSVPDHGH 578
QY 825 EGL--LLGFPQKAYAPFLHWTGRVVLFLFALFGVSLYSCHISVGLDQELALPKDSY 882
Db 579 TDHPISLFRDYFGFPLTRSESKYFVVIYLYIISIIYGGCFHQBEGDLRLNLASDDSY 638
QY 883 LLDYFLFLNRYF-EVGAPVYFVTTLGYNFSSEAGNMAICSSAGCNFSTQKIYATEPP 941
Db 639 ITPYFNVEENYFSDYGRVMVITKVDYWDK-----DVRQKLENTKIP 683
QY 942 EQSYLAIP--ASSWVDDFDMLTPSSCCRLYISGPNKDKFCPSTVNSLNCNCHSITMG 999
Db 684 EKNVYVDKLTPEFLDAVQ-----YLGNSQD---PNEKNT-----717
QY 1000 SVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVNLTSDGVLASR 1048
Db 718 -----FMNIPDLSNFPN-----FQHDINISSNEIISR 748

RESULT 71
US-10-094-749-1772
; Sequence 1772, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: TSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
```

Db 550 ETNREELYDLLETLREL--SVTSKVKFIV-FNPSP-----VTMDRYASSL--GAPL 595

Qy 1116 LSLCLVPTTAV--SCLLGLDRLSGLNL--LSIVMILVDTVGFMAWDISNAVSLIN 1170

Db 596 HNSCISALFLFPFSAFLVA---DSLINWITLTVVSVEFGVIGFWTLWKVELDCISVLC 651

Qy 1171 LVSAGVMSVEFVSHITRSPAIS---TKPTWLERAKEATISMSGSAVFAGVAMTN-----LP 1222

Db 652 LIYGINVITDNCAPMSAFVLGKDFTRTKWKNALV-----HGVAILOSYLCYIV 702

Qy 1223 GILVGLAKAQLIQIPFFRLNLITLLGLHLVFLPVLISYVGPVDPNPALEAEOKEAEE 1282

Db 703 GLIPLAAVPSNL-TCTFLRCLFLIAFTVFFHCFAILPVILTFLPPS-----KCKRKEK 754

RESULT 72

US-10-060-756A-10

; Sequence 10, Application US/10060756A

; Publication No. US20030046717A1

; GENERAL INFORMATION:

; APPLICANT: Zhang, Jian

; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN

; FILE REFERENCE: PB0177

; CURRENT APPLICATION NUMBER: US/10/060,756A

; CURRENT FILING DATE: 2002-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00667

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00664

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00669

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00665

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00668

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: US 09/864,761

; PRIOR FILING DATE: 2001-05-23

; PRIOR APPLICATION NUMBER: US 60/327,898

; PRIOR FILING DATE: 2001-10-09

; NUMBER OF SEQ ID NOS: 4804

; SOFTWARE: Acomica Sequence Listing Engine

; SEQ ID NO 10

; LENGTH: 648

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-060-756A-10

Query Match 4.5%; Score 310.5; DB 4; Length 648;

Best Local Similarity 22.0%; Pred. No. 2.1e-17;

Matches 133; Conservative 107; Mismatches 238; Indels 127; Gaps 24;

Qy 334 HT-----LLGQFFQCGMTWVASWPLTILVSVIPVVALAAGLVF--TELTTDPVELWSAP 386

Db 115 HTDCLGSLLSRTFQWLGWQVGAHPWIFLLAPLMTAALGTLGTLGFLYLPKDEEDLEEHYTPV 174

Qy 387 NSQARSEKAFHDOHFGFPFRTNQVILTAPNRSS-----YRYDSLILGPKNFGSIL 436

Db 175 GSPAKAERRFVOGH-----FTTNDYSRFSASRRSTEANFVSLLVVSYSDSLDDPATFAEVS 230

Qy 437 DLDLLLELLEQLRHLRHLQVSPCAORNISLQDIC--YAPLNPNDTSLYDCCI--NSLLQ 492

Db 231 KLQGAQVDLAV-----AREKSGIQYQQVCARYAL-----CVPPNPILY 270

Qy 493 YFQNNRTLLLTANQTLMGOTSOVDWKDHPFLCANAPLTFKDGTLALSCWADYGA-PVF 551

Db 271 AMQVNTLNL-----SSISFPAY-----NNGRHPLY 296

Qy 552 PFLAIGY-----KGDYSEAEALMTFSLNNYPAGDPRL-AQAKLWEEAFLEEMRAF 603

Db 297 LTGPFPGYILGSGLMGQLLLRKAMRLLYLK---TEDPEYDVQSKQWLTHLLDQFTNI 353

APPLICANT: IRIE, RYOTARO

APPLICANT: TAMECHIKA, ICHIRO

APPLICANT: SEKI, NACHIKO

APPLICANT: YOSHIKAWA, TSUTOMU

APPLICANT: OTSUKA, MOTOTUKI

APPLICANT: NAGAHARI, KENJI

APPLICANT: MASUHO, YASUHIKO

TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA

FILE REFERENCE: 084335/0160

CURRENT APPLICATION NUMBER: US/10/094,749

CURRENT FILING DATE: 2002-03-12

PRIOR APPLICATION NUMBER: 60/350,435

PRIOR FILING DATE: 2002-01-24

PRIOR APPLICATION NUMBER: JP 2001-328381

PRIOR FILING DATE: 2001-09-14

NUMBER OF SEQ ID NOS: 3381

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 1772

LENGTH: 783

TYPE: PRT

ORGANISM: Homo sapiens

US-10-094-749-1772

Query Match 4.9%; Score 339; DB 4; Length 783;

Best Local Similarity 21.1%; Pred. No. 9.3e-20;

Matches 190; Conservative 161; Mismatches 333; Indels 216; Gaps 41;

Qy 444 LLEQLERHLQVWPEAQRNLSLQDICYAPLNPNTSLYDCCINSLLQYFON----NRT 499

Db 10 ILKHAADVTKIQV--PRPGNFYTFAHICI--LNDKTKTIVDDIVH-VLEELKNARATWRT 64

Qy 500 LLLLTANQTLMGTSQVDKDHFLYCANAPLT-FKDGTLALSCWADYGAVPFPLAIG 558

Db 65 NFAIT-----YPIHLKDGRA-----VYNGHQLGG 89

Qy 559 --YKGD-YSEABALIMTFLNNYPAGDPRLAQAQWEEAFLEEMRAFQRRMAGFQVTF 615

Db 90 VTVHSKDRVKSAAEQITVYLOSINSLNDWAER--WESSFCDTVALFQKSNKVKWYYP 147

Qy 616 TARSLEDEINRTAEDLPFATSYIVILYISLALGYSNRSRVWVDSKATLGLGGVAV 675

Db 148 TSS-SLREDPQKTSRVSERYLVTSILV---VTMAILCCS--MDCVRSKPWLGLGLVT 201

Qy 676 VLGAVMAAGFFSYLGRSLVTLQVVPFLVLSVGADNIFIVLEYQRLPRPGEPREVH 735

Db 202 ISLATLTAGIINLTGKYNSTFLG-VPFVWLGHGLYGTFFEMLSM---RKTREDQHVK 256

Qy 736 IGRALGRVAPSMLLCSLSEA---ICPFLGA--LTPMPAVRTFALTSGLAVIDFLQMSA 790

Db 257 -ERTAAYVADSMLSFSLTTAMYLVTFGIGASPTNIEAARIFCCNSCIAIFFNYLYVLSF 315

Qy 791 FVALLSLSKQOASRLDVCCVKP-----QELP-----PPQQGS----- 825

Db 316 YGSSLVFTGYENNYQHSIFCRKVPKPEALQEKPAWTRFLLTARFSDTAEGEANTYES 375

Qy 826 GLLLGFFQKAPAPFLHWRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLD 885

Db 376 HLLVCFKRYKRYCDWITVYKPFVFLYIVISFALMGYLQVSGSLSNIVATATOTIE 435

Qy 886 YFLFLNRYFEVGAIV---YFVTLTGYNFSS-----EAGNAICSSAGCANNFSTQKIY 936

Db 436 YTTAQKQYFSNYSPIGVIGYIESIEYWNYSQVEDVLEYTKGFVIRISFESYLNLRKLVN 495

Qy 937 ATEPEQSYLAIPASSWDDFDMLTPSSCCRLVYISGNKDKFCPTVNSLNCILKNOWSI 996

Db 496 STGLPKK-----NFTDMLRNS-----FLKAPQSFHQEDIFIS----- 528

Qy 997 TWGSRVPSVEQPHKYPWFLNDRPNIKCPKGLAAYSTSVNLTSDGQVLAASFWMYHKPL 1056

Db 529 -----KKY-----ND-----EVDVVASRMFLVAKTM 549

Qy 1057 K-NSQDYTEALRAARELANITADLRKVPDGPDPAPFEPYITVNFYEQYLITLPEGLFM 1115


```
QY 604 QRMMA--GMQVTF--ERSLEDEINRTAEDLPIFATSVIIVFLYSLALGSYSWSRV 660
DB 354 KNILALKKIEVHFTSLRQLEFEATSVTV--IPVPHLAYILLILFAVTSCTFRFD-----406
QY 661 MVDKATLGLGGVAVVILGAVMAAGPFSYLGIRSSILVILQVPPFLVLSVAGADNIFIVLE 720
DB 407 CIRNKCVAAGFVISAFVAVSGFGLLHIGV--PFVLIIVANSFPLILGVGVDDMFIMISA 465
QY 721 YQRLPRRGPBREVHIGRALGRVAPSMMLCSLSEACIFFLGALTPMPAVRTALTSLGLAV 780
DB 466 WHK--TNLAGDIRE--RMSNVYSKAAVSITITITINILALYTGIMSSFRSVCQCIYTGTTL 523
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKVQOELP-----PPGQG 824
DB 524 LFCYFNITCFGAFMALDGREV-----VCLWLKADPKWPSFKKCFCCFPFGSVDEHG 578
QY 825 EGL--LLGFPQKAYAPFLHWTIRGVVLLFLALFGVSLYSMSCHISVGLDQELALPKDSY 882
DB 579 TDIHPSLFRDYFGPFLTRSESKYFVVIYVLYIISIIYGCFFHVQEGDLRLNLASDDSY 638
QY 883 LLDYF 887
DB 639 ITPYF 643

RESULT 73
US-10-890-776A-10
; Sequence 10, Application US/10890776A
; Publication No. US20050129683A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/890,776A
; CURRENT FILING DATE: 2004-07-14
; PRIOR FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 10
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-890-776A-10

Query Match 4.5%; Score 310.5; DB 5; Length 648;
Best Local Similarity 22.0%; Pred. No. 2.1e-17;
Matches 133; Conservative 107; Mismatches 238; Indels 127; Gaps 24;

QY 334 HT-----LLGQFPQGGTWWASWPLTILVLSVIVVLAAGLVF--TELTTPDVELMSAP 386
DB 115 HTDCLGELLSRTFQWLGMQVGAHPWIFLLAPLMLTAALGTGFLYLPKDEEDLEEHYTPV 174
QY 387 NSQARSEKAFDQHFQFPFTNQVILTAPNRSS-----YRDSLLLGPKFSGTL 436
DB 175 GSPAKAERRFVQGH-----FTTNDYRFSASRRSTRANFVSLVLSVSDSLDLPATFAEVS 230
```

```
QY 437 DLDLLELELELRLHQLQVMSPEAQRNLSLQDIC--YAPLNPDTNLSYDCCI--NSLLQ 492
DB 231 KLDGAVQDLRV-----AREKGSQIQYQVCARYRAL-----CVPNPILY 270
QY 493 YQNRNETALLTANQTLMGQTSQVDMKDHFLYICANAPLTFKQGTALALSCHADYGA--PVF 551
DB 271 AMQVNTKLNL-----SSISFPAY-----NHGRHPLY 296
QY 552 PFLAIGY-----KGKDYSEAEALIMTFSLNNYPAGDPRLL--AQAKLWEEAFLEEMRAF 603
DB 297 LTGFFGYILGSLGQGLLRKANRLIYLYLK---TEDPEYDVQSKQWLTHLLDQFTNI 353
QY 604 QRMMA--GMQVTF--ERSLEDEINRTAEDLPIFATSVIIVFLYSLALGSYSWSRV 660
DB 354 KNILALKKIEVHFTSLRQLEFEATSVTV--IPVPHLAYILLILFAVTSCTFRFD-----406
QY 661 MVDKATLGLGGVAVVILGAVMAAGPFSYLGIRSSILVILQVPPFLVLSVAGADNIFIVLE 720
DB 407 CIRNKCVAAGFVISAFVAVSGFGLLHIGV--PFVLIIVANSFPLILGVGVDDMFIMISA 465
QY 721 YQRLPRRGPBREVHIGRALGRVAPSMMLCSLSEACIFFLGALTPMPAVRTALTSLGLAV 780
DB 466 WHK--TNLAGDIRE--RMSNVYSKAAVSITITITINILALYTGIMSSFRSVCQCIYTGTTL 523
QY 781 ILDFLLQMSAFVALLSLDSKQASRLDVCCCKVQOELP-----PPGQG 824
DB 524 LFCYFNITCFGAFMALDGREV-----VCLWLKADPKWPSFKKCFCCFPFGSVDEHG 578
QY 825 EGL--LLGFPQKAYAPFLHWTIRGVVLLFLALFGVSLYSMSCHISVGLDQELALPKDSY 882
DB 579 TDIHPSLFRDYFGPFLTRSESKYFVVIYVLYIISIIYGCFFHVQEGDLRLNLASDDSY 638
QY 883 LLDYF 887
DB 639 ITPYF 643

RESULT 74
US-10-060-756A-4799
; Sequence 4799, Application US/10060756A
; Publication No. US20030046717A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/060,756A
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 4799
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-756A-4799

Query Match 4.4%; Score 306.5; DB 4; Length 648;
Best Local Similarity 21.8%; Pred. No. 4.7e-17;
```

```

; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4809
; SOFTWARE: Aecomica Sequence Liating Engine
; SEQ ID NO 4799
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-890-776A-4799

Query Match      4.4%; Score 306.5; DB 5; Length 648;
Best Local Similarity 21.8%; Pred. No. 4.7e-17;
Matches 132; Conservative 108; Mismatches 238; Indels 127; Gaps 24;

QY 334 HT-----LLGQPPQGGTGWVASHVPLILVLSVTVVALAAGLVF--TELITDPVELKSP 386
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 HTDCLBGLLSRTQWLQGWQGAHPWIFLLAPLMATAALGTGFLYLPKDEEDLEHPIPV 174
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 387 NSOAREKAFHDQHFQFPFRNTQVILTAPNRSS-----YRDSLLIGPNFSGIL 436
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 GSPAKAERRVQGH-----FTNDYSYFASERSTEANFVSLVVSYSDSLDPATFAEVS 230
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 437 DLDLLLELLELOBLRHLQWSPEAQRNISLQDIC--YAPINPDNTSLYDCCI--NSLLQ 492
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 KLDGAVQDLRV-----AREKSGIQVQVCARYAL-----CVPPNPFLY 270
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 493 YFQNNRTLLLTANQTLMGQTSQVDWDXHFLYCANAPLTFKDGATALASCMWYGA-PVF 551
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 271 AWQVNTKTLN-----SSISFPAY-----NHGRHPLY 296
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 552 PELAIQGY-----KGKDYSEAEALMTFTSLNNYPAGDPLR-AQAKLWEBAFLUEEMRAF 603
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 LTGFFGGYILGSGLGMCQALLRAKAMLLYYLK--TEDPEYDVQSKQWLTHLLDQFTNI 353
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 604 QRRMA--GMFQVTTA-ERSLEDEINTTAEIDLPIFATSYIVIFLYISLALGSYSWSRV 660
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 KNILAKKIEVHPHTSLSRQLEPATSVT--IPVFLAYILILFVATSCFRPD-----406
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 661 MVDSEKATLGLGGVAVLVGAVMAAGFFSYLGRISLVILQVVPFLVLSVGADNIPIFVLE 720
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 407 CIRNMCMVAAPGVLSAFNAVVSFGFLLHIGV-PPVILIVANSPFLILGVGVDDMFIMISA 465
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 721 YQRLPRRPGEPREVIHGALGRVAPSMLLCSLSEAI CFFLGALTMPAVRTTALTSGLV 780
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 466 WHK-TNLADDIRE-RMSNVYSKAAVSIITITITNIALYTGIMSSFRSQVCFCIYGMTL 523
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 781 ILDFLLQMSAFVALLSLDSKQEASRLDVCCKVKQELP-----PPGQG 824
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 524 LFCYFYNIITCGAFNALDGKREV-----VLCWLKAKPKWPSFKKFCFPFGSVDEHG 578
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 825 EGL--LLGFFQKAVAPFLLHMTITRGVWILLFLALFGVSLYSMCHISVGLDQELALPKDSY 882
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 579 TDIHPISLFFRDYFGPFLTRSESKYFVVFVLYVLYIISIIYGCFFHQVEGLDLRLNASDYS 638
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 883 LLDYF 887
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 639 ITPYF 643
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 76
US-09-864-761-34826
; Sequence 34826, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES US
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23


```
QY 1274 ALBOKRABEA 1283
Db 118 PIEQSQPPSA 127

RESULT 79
US-10-415-934-8
; Sequence 8, Application US/10415934
; Publication No. US20040030099A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN
; FILE REFERENCE: LIO204 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/415,934
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US 60/245,565
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: US 60/245,564
; PRIOR FILING DATE: 2000-11-06
; PRIOR APPLICATION NUMBER: US 60/245,572
; PRIOR FILING DATE: 2000-11-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 792
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-415-934-8

Query Match 3.9%; Score 266.5; DB 4; Length 792;
Best Local Similarity 21.1%; Pred. No. 28-13;
Matches 140; Conservative 116; Mismatches 259; Indels 149; Gaps 26;

QY 687 PSYLGIRSLVLQVVPFLVL-SVGADNIFIFVLEYQRLPRPGEPRVHIGRALGRVAP 745
Db 227 PEFDSLESVLIVKAVHNLQSHGKGVFELSGWR--TKENLPFKDRIADAYSDVMV 284

QY 746 SMLLCSLSBAICFFIGA--LTPMPAVRTFALSGLAVIDFLQWSAFVALLSLQKQE 803
Db 285 TYTMTSSLYFITFGMGASPTNIEAVKFCQNMCSILLNLYFIYFISFGSCLVFAGQLEQ 344

QY 804 ASRLDVCCCVK-----QELPPGGE-----GLLGFQKAYAPL 840
Db 345 NRYHSIFCCKIPSAEYLDKRVWFQVMSDGHQOOTSHTNETHYQHFIQHLEHNEWI 404

QY 841 LEWITRGVLLFLALFGVLSYMSCHISVGLQELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 405 THIVYKPFVILYLIYASFSGCCLQISDGANIINLLASDSPSVSYAMVQKYFNSYSPV 464

QY 901 ----YFVTTLGY-NFSSEAGMNAICSSAGCNFTSFTOKIQYATEFPFQSYLAIPASSWDD 956
Db 465 IGFYVVEPLEYWNSSVQDLRLCS-----GFT-----AVSWVEQ 499

QY 957 FIDWLTSPSCCRLYISGPNKDKPCPTVNSLNCNCSITMGSV--RPSVQPFKYLWP 1014
Db 500 YYQFLKVS-----VSANKSDP-----ISVLOSSFLUKPEFQHF----- 534

QY 1015 FLNDRNIPKPGGLAAYSTSVNLTSQGVLSRFMAYHKPLKNSQ-DYTEALRAARELA 1073
Db 535 -----RNDIIFSKAG-----DESNIIASRLYLVARTSDKQKEITEVLEKRLPL- 578

QY 1074 ANITADLRKVPGDPAPEFVFYITNVFVEQY-----LTLPEGLFMLSCLVPTTAVSC 1128
Db 579 -SLSKSIRFIV-FNPSF-----VFMDHYSLSVTVPVLIAG-FGVLLVLILITFELVI 626

QY 1129 LLLGLDLRGLNLSIVMLVDTVGFMAIWDISYNVSLINLVSAVGSVFVSHITS 1188
Db 627 HPLG-----NFWLILSVTSIELVGLMLTWNVDMDCISILCLITYTINFAIDHCAPLLFT 691

QY 1189 FAISTKPTMLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQIQLIQLIFFFLNLLITL 1248
Db 682 FVLATEHTRTQCICKSSLQDHGTAL-----LQNTVSFLI-----GLVPLLFPVSNLFTL 730
```

```
QY 1249 L-----GLLHGLVFLPVILSYGVGPDVNPALALEOKRAE-----EAVAAVMVASCPNH 1295
Db 731 FKCLLLTGGCTLLHCFVILPVFLTFPPP--SKKHKKKKRAKREKEREIEICIEIQENFDH 788

QY 1296 PSRV 1299
Db 789 VTTV 792

RESULT 80
US-10-424-599-181755
; Sequence 181755, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: SOY Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181755
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(229)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135138C.1.pap
US-10-424-599-181755

Query Match 3.8%; Score 264.5; DB 4; Length 229;
Best Local Similarity 28.4%; Pred. No. 4.3e-14;
Matches 75; Conservative 44; Mismatches 102; Indels 43; Gaps 8;

QY 369 GLVFTLTTDPVELMSAPNSQARSEKAFHDQHFGFPFRTNQVILTPAPNRSYRYSLLLG 428
Db 4 GLLREFAEETREPKLVWPGSKAAEKEFFDLSHLPFYRIEQILTIATIPESKHGKPPSIIT 63

QY 429 PKNFGIIDL DLLLELLEQLERLHLQVWSPAQRNISLQDICYAPLNPDNTSLYDCCIN 488
Db 64 EEN-----IELFIEIQEKVDGIR-----ANYSGLLVSLSDICLPLGD-----DCASQ 106

QY 489 SILQYFQNNRTLLLTANQTLMGQTSQVDWKDHPLYCANAPLTFKDGDTALALSCMADYGA 548
Db 107 SILQYFQ-----MDPDNYDXLRXCXHAERYWFQHYTSTE-TCFSAFKA 147

QY 549 PVFPELAIGYKGDYSEAEALIMTFSLNN--YPAGDPR---LAQAKLWEEAFLEMRAP 603
Db 148 PLEPTTALGXSGNNYSEASAFVITVPVNNATKVGDENGKAIXGKAFIQLAKEBLLPM 207

QY 604 QRRMAGMFQVTTAERSLEDEINR 627
Db 208 VQ--SNNLTSPSTESSIEELKR 229

RESULT 81
US-10-425-115-201210
; Sequence 201210, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 201210
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(267)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_11508C.1.pep
US-10-425-115-201210

Query Match      3.7%; Score 259; DB 4; Length 267;
Best Local Similarity 29.5%; Pred. No. 1.6e-13;
Matches 77; Conservative 38; Mismatches 108; Indels 38; Gaps 13;

QY 20 QSEPYTHIQPGYCAFDECGKPELGSMTLSVSCLSNTPARKITGDHLI--LLQKI 77
DB 34 QNSGSRVPAEGYCSMYGICAQRSDD--GKVLNCANAT-----KAVKPDTLFSSRIQSL 84
QY 78 CPRLYTGPNQACCSAKQLVLSLASLITKALLTRCPACSDNFVNLHCHNTCSPNQLFI 137
DB 85 CPPI--TG--DVCTVDQFDTLHQVQVAVPFLVGCPCACLRNLFNFCMSCSPNQLFI 140
QY 138 NVTRVQLGAGQLPAAVVAEAFYQHGFPAEQSYDCSRVRPAAATLAVGTMCQGVGSALC 197
DB 141 NTVSVKQINSTM--TWDGIDYVYTHYGBELVNSCKDVKFGTINTAMPFI-----GAGAK 194
QY 198 NQQRWLNFP---QSDTNGNGLAPLDTITFHLLEPGQ--AVGSGIQPLNGBVARCNESQGDVVA 252
DB 195 TYKDWLAFTXROANLNEPSPYLITF-----GSDPDSGSKVFLXSTIYSC-----CTPFL 245
QY 253 TCS-CODCAAS--CPAIRPQ 270
DB 246 XCSXCGDCPSSXKAGSLLPQ 266
```

```
RESULT 82
US-10-408-765A-1316
; Sequence 1316, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Wainock, Dale B.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1316
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-1316

Query Match      3.6%; Score 252; DB 4; Length 505;
Best Local Similarity 20.9%; Pred. No. 1.8e-12;
Matches 124; Conservative 105; Mismatches 219; Indels 146; Gaps 24;

QY 756 ICFELGA--LTPMPAVRTALTSLGLAVILDFLLQMSAFVALLSLDSKQSESLDVCCCV 813
DB 8 ITFGMGASPTTIEAVKVPQNMVCSILLNYFIYFSPGSCLVFAGQLEQNRHYSFCK 67
```

```
QY 814 KP-----QELPPPGGGE-----GLLLGFFQKAYAPFLLHMTIRGWL 850
DB 68 IPSAEYLDRKPVWFQTVMSDGHQTSHTHETNPQHHPHQHFLREHYNEWITNIYVKFPV 127
QY 851 LLEFALFGVSLYSMCHISVGLDQELALPKDSYLLDYPLFLNRYFEVGAPV---YFVTTLG 907
DB 128 ILYLIYASFSFMGCLQISDGANIINLLASDPSVSAMVQOKYFSNYSPIGVFYVPEPLE 187
QY 908 Y-NFSSEAGMNAICSSAGCNPSPSTQKIQVATFEPPQSYLEIAIPASSWVDPDFDWTSSC 966
DB 188 YWSSVQDDLRRLCS-----GFT-----AVSWVEQYQFLKVSN- 221
QY 967 CRLYISGPNKDKPCPSTVNSLNCNKMCSITMGSV--RPSVEQFHKYLPWFLNDRPNIK 1024
DB 222 ----VSANKNSDF-----ISVLQSSFLKKPEFQHF-----RNDIIF 253
QY 1025 PKGGLAAYSTSVNLTSDGQVLAIRFMAYHKPKNSQ-DYTEALRAARELANITADLRKV 1083
DB 254 SKAG-----DESNIIASRLYLVA RTSRDKQKEITEVLEKRLPL--SLSKSIRFI 300
QY 1084 PGTDPAFEPVFPYITNVFYEQY-----LTILPEGLFMLSICLVPTTAVSCLLGLDLRS 1138
DB 301 V-FNPSF-----VFMDHYSLSVTVPVLIAG-FGVLLVLITLFTFLVIHPLG-----N 344
QY 1139 LMLLSIVMLVDTVGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWL 1198
DB 345 FWLILSVTSIBELGVGLMTLWVDMDCISILCLIIYTNFAIDHACAPLLFTFVLATEHRT 404
QY 1199 ERAKEATISMGSAVFAVAMTNLPGLIVLGLAKAQIQLIPFRLNLLITLL----- 1249
DB 405 QCIKSSLDQHGTAI-----LQNVTSPLI-----GLVPLLFPVPSNLTFTLPKCLLLTGGC 453
QY 1250 GLLHGLVLPVLISVGVDPVNPALALEQKRAE-----EAVAAMVAVSCPNNHPSRV 1299
DB 454 TLLHCFVILPVFLTPPPP--SKKHKKKRKRKREREIEIEIQENPDHVTTV 505

RESULT 83
US-10-432-613-2
; Sequence 2, Application US/10432613
; Publication No. US20040048282A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN PATCHED-LIKE PROTEIN
; FILE REFERENCE: L10234 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/432,613
; CURRENT FILING DATE: 2003-06-04
; PRIOR APPLICATION NUMBER: US 60/251,032
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-432-613-2

Query Match      3.6%; Score 245.5; DB 4; Length 1491;
Best Local Similarity 20.6%; Pred. No. 3.6e-11;
Matches 247; Conservative 150; Mismatches 494; Indels 307; Gaps 52;

QY 213 LAPLDITP-HLLEPGQAVSGIQPLNEGVARCNESQGDVATCSQDCAASCFAIARPOA 271
DB 179 LAPAHFTYPRALQYQG-GSSLPGL-----GDRAALCSHGSSLSPSPA---PSQ 223
QY 272 LDSTF-----YLGQMPGS-----LVLIILCSVFAVVTILLVGF 306
DB 224 RGTWKPPAVQHHVSVRQERAFQMPKSVSQAIAEMPVAVLMLCLAVIPLCTLAGLGR 283
QY 307 VAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFPQGMGTWVASWPLT-----ILVLSVTP 362
DB 284 L-PDFSFKPLLAGPEPRDITDIGSKLV-----VMRALQALTQPRKLLFLS--P 325
```


Db 658 GGIVHTRILVERLTCLLTSTVLAIVNLKVALMISISPPQLRPSSELL-----TQ 704

Qy 1024 CPRGGLAAYS-----TSVNLTSQGVQLASRFMAYHKPLKNSQDYTE 1064

Db 705 GPRGTLAHISQLKSLSPSPFANFCLTADSKARSQLL-----PSGNCPIYAE 753

RESULT 85

US-10-017-161-2024

; Sequence 2024, Application US/10017161

; Publication No. US20030143668A1

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 084335/0152

; CURRENT APPLICATION NUMBER: US/10/017,161

; PRIOR FILING DATE: 2002-12-18

; PRIOR APPLICATION NUMBER: JP 2001/246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2430

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2024

; LENGTH: 542

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-017-161-2024

Query Match

Best Local Similarity 3.3%; Score 225; DB 4; Length 542;

Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;

Qy 744 APSMLCSLSEA---ICFPLGA--LTPMPAVRTFALTSLGLAVILDPLLOMSAFVALLSLD 798

Db 23 AUSLSFSLTAMLYVTFGIGASPTNIEARIFCCNSCIAIFNFNLYLSFGSSLVFT 82

Qy 799 SKRQASRLDVCCVKP-----QELP-----PPGQGE-----GLLLGFFQ 833

Db 83 GVIENNYQHSIFCRKVPKEALQEPAWYRFLLTARFSDTAGGEANTYVESHLLVCFLK 142

Qy 834 KAYAPFLHWTIRGVVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893

Db 143 RYCDWITNTYVKPFVFLYIISPALMGYLQVSGSDLSNIVATATQIETTTAAQKY 202

Qy 894 FEVGAPVYFVTTLGYNFSEAGNAICSSAGCANNFSTQKIQVATFPPSQSYLAIPAS-- 951

Db 203 FSNYSPVI-----GFYIESIEYWNITSVQEDVLEYTKGFV 237

Qy 952 --SWDDFDLMTTPSSCCRLYIS--GPNKDKFCPSTVNSLNCNKCNSITMGSVRPSVEQF 1008

Db 238 RISWPSYLYNLR-----KLVSTGLPKGNFTDMLRNSP--LK-----273

Qy 1009 HKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVQLASRFMAYHKPLK--NSQDYTEALR 1067

Db 274 ---APQSFHQEDIIIPSK-----KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE 320

Qy 1068 AARELAANITADLRKVPGTDPAPFVPPYITNTVFYEQYLTILPEGLFMLSCLVPTFAV- 1126

Db 321 TLRLR--SVTSKVKFIV-FNPSF-----VYMDRYASSL--GAPLHNSCISALFLFLF 366

Qy 1127 -SCLLLGLDLRSLGLNL---LSIVMLVDVTGFMALWDISYNVAVSLINLVSAGMSVRFV 1182

Db 367 FSAFLVA---DSLINVMTLTVSVFVGIGFMTLWKVELDCISVLCLYIYINITYDNC 422

Qy 1183 SHITRSPAIS---TKPTWLERAKEATISMGSAVFAGVAMTN-----LPGILVLGLAKAQL 1234

Db 423 APMLSTFVLGKDFTRTKWKVNALEV-----HGVALQSYLCVIVGLIPLAAVPSNL 473

Qy 1235 IOIFFPRNLNLTLLGLLHGLVFLPVILSYVGPVDVNPALAEQKRAE 1282

Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKGRKEK 513

Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKGRKEK 513

RESULT 86

US-10-292-798-1670

; Sequence 1670, Application US/10292798

; Publication No. US20030235833A1

; GENERAL INFORMATION:

; APPLICANT: SUWA, MAKIKO

; APPLICANT: ASAI, KIYOSHI

; APPLICANT: AKIYAMA, YUTAKA

; APPLICANT: ABURATANI, HIROYUKI

; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS

; FILE REFERENCE: 084335/166

; CURRENT APPLICATION NUMBER: US/10/292,798

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: 10/017,161

; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: JP 2001-246789

; PRIOR FILING DATE: 2001-06-18

; NUMBER OF SEQ ID NOS: 2070

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1670

; LENGTH: 542

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-292-798-1670

Query Match

Best Local Similarity 3.3%; Score 225; DB 4; Length 542;

Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;

Qy 744 APSMLCSLSEA---ICFPLGA--LTPMPAVRTFALTSLGLAVILDPLLOMSAFVALLSLD 798

Db 23 AUSLSFSLTAMLYVTFGIGASPTNIEARIFCCNSCIAIFNFNLYLSFGSSLVFT 82

Qy 799 SKRQASRLDVCCVKP-----QELP-----PPGQGE-----GLLLGFFQ 833

Db 83 GVIENNYQHSIFCRKVPKEALQEPAWYRFLLTARFSDTAGGEANTYVESHLLVCFLK 142

Qy 834 KAYAPFLHWTIRGVVLLFLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRY 893

Db 143 RYCDWITNTYVKPFVFLYIISPALMGYLQVSGSDLSNIVATATQIETTTAAQKY 202

Qy 894 FEVGAPVYFVTTLGYNFSEAGNAICSSAGCANNFSTQKIQVATFPPSQSYLAIPAS-- 951

Db 203 FSNYSPVI-----GFYIESIEYWNITSVQEDVLEYTKGFV 237

Qy 952 --SWDDFDLMTTPSSCCRLYIS--GPNKDKFCPSTVNSLNCNKCNSITMGSVRPSVEQF 1008

Db 238 RISWPSYLYNLR-----KLVSTGLPKGNFTDMLRNSP--LK-----273

Qy 1009 HKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVQLASRFMAYHKPLK--NSQDYTEALR 1067

Db 274 ---APQSFHQEDIIIPSK-----KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE 320

Qy 1068 AARELAANITADLRKVPGTDPAPFVPPYITNTVFYEQYLTILPEGLFMLSCLVPTFAV- 1126

Db 321 TLRLR--SVTSKVKFIV-FNPSF-----VYMDRYASSL--GAPLHNSCISALFLFLF 366

Qy 1127 -SCLLLGLDLRSLGLNL---LSIVMLVDVTGFMALWDISYNVAVSLINLVSAGMSVRFV 1182

Db 367 FSAFLVA---DSLINVMTLTVSVFVGIGFMTLWKVELDCISVLCLYIYINITYDNC 422

Qy 1183 SHITRSPAIS---TKPTWLERAKEATISMGSAVFAGVAMTN-----LPGILVLGLAKAQL 1234

Db 423 APMLSTFVLGKDFTRTKWKVNALEV-----HGVALQSYLCVIVGLIPLAAVPSNL 473

Qy 1235 IOIFFPRNLNLTLLGLLHGLVFLPVILSYVGPVDVNPALAEQKRAE 1282

Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFPLPS-----KKGRKEK 513

Db 882 GNEAAYSTNGSTFTPLLVKALVIEFESNVAISTYIYANIRQFYSEVHEWFMQOLKTAPPEL 941
Qy 1073 AAN-ITADLRKVPGTDPFAFEVFPYITITNVFYEQYLTLPEGLPMLSLCLVPTFAVSCLLL 1131
Db 942 QCGWFTSDLK-----FYNVQDTLSHDTFVAICLAWAASLAV-LLCF 981
Qy 1132 GLDLRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVFEVSHITRSPAI 1191
Db 982 TVNILISIVAVLTVSISIFNTVAVLILGWQLNILESIAVSTAIGLAVDFSLHYGIHYRM 1041
Qy 1192 STKPTWLERAKEATIS--MGSAPVAGVAMTNLPGILVGLAKAQLIQIFFPRLNLLITLL 1249
Db 1042 SPVKERL-AATQFVLSRIIGPTVMAATT-TGLAGGIMMASNLPIYIQIGVFL--VVVMIV 1097
Qy 1250 GLLHGLVFLPVILSYVGPD-----VNPALALEQKRAEAV 1284
Db 1098 SWFYATFLLMSLLRVAGPQHGLEKWLKWSKSSGSKFYERKPSQVIASEQLLTPSS 1157
Qy 1285 AAWVASCNHPRSVSTADNIYVNHSPFGSIKAGAI SNFLPN 1328
Db 1158 AIVELANSETHELESINSNLIKTIIS---GIESAHALSS-LPRD 1197

RESULT 89

US-11-097-143-3123
; Sequence 3123, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3123
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: DROSOPHILA

US-11-097-143-3123
Query Match 3.1%; Score 214.5; DB 6; Length 1218;
Best Local Similarity 18.7%; Pred. No. 1.3e-08;
Matches 210; Conservative 185; Mismatches 412; Indels 317; Gaps 50;

Qy 374 ELTDPVELWSPNQAQSEKAFHQHGFPPRTNQVILTAPNRSSRYDSLLLGKPNFS 433

Db 222 CAATLFTDGMHRRQPTIEGFFCD-----SSPRKEYSHFVVQRIQGNATD 267

Qy 434 GILDLDLLELELOERLHLQVWSPQAQNSLQDIC-----YAPLNPNTSLYDCC 486

Db 268 SLFDLNGLLAMCOLQDQITEVPSYRACFEPMLTTECCRPWSPNPAALANKSCFDD-- 325

RESULT 90

US-10-168-428-2
; Sequence 2, Application US/10168428
; Publication No. US20030165897A1

Qy 487 INSLQYFQNNRTLLLTANQTLMGQTSQVDMKDH-----LYCANAPLTPKDGATLAL 540
Db 326 -----LTTEDEVSTLHTLLGCVYEPHDLKMDNHCNEIP----- 358
Qy 541 SCHADYGAVPFPLAIGGKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
Db 359 HCRAP-----BECKRLNIVFNLNF-----LTDSPISKSN 388
Qy 601 R-----AFORMMAGMFOVTFATERSLEDEINRUTTABDL----- 633
Db 389 DSNVLYKAMIPIVPAQSNELLPLFHEWEDVZ--LINELVEVVAMDGLLENELFNBLLLT 446
Qy 634 PIPATSYIVIFLYISIALGSYSWSKVMVDSKATGLGGVAVVLGAVMAAMGPFYSLGR 693
Db 447 DVMVLSLGGTFVMAWSVWLYTGSATITLMSCAVAFSL-GLAVFFYAI VLEPFPFYMNLL 505
Qy 694 SSVILQVVPFVLVLSVGADNIPFVLEY-----ORLPRR-----PGH 730
Db 506 AVVVI-----IGIGADDDVFLFKIMHCVLTERPFSNRCITLTQSQSALPTLENSDHT 557
Qy 731 PREVHIGRALGRVAPSMILCSLSEAI CFFPLGALTMPAVRTALTSGLAVIDLFLQMSA 790
Db 558 SLEIMALTMRHAAASMFVSTLTAGAFYASVSSSITAIKCGIFAGTVVNYLLMITW 617
Qy 791 FVALLSLDSKROEASRLDVCCVFPQELPPQGGELLAGFFQKA---YAPFLHWHITRG 847
Db 618 LPASVSI-MERLFAITMS--CHHPMSIK-----LIHACKKSINRCQMPESCITKS 665
Qy 848 VV-LLLFLALFGVSLYSMCHISVGLDQELALPKOSYLLDYDFLFLNRY-FEYGAPV----- 900
Db 666 IMNYAYLWLLIFG-ALGASSAVIVFWYFGLQLPKESH---FQLEFVSKHPFEVYSSLKQOF 721
Qy 901 YPVVTTL-GV-NFSSSEA---GMNAICSSAGCNPFSFTQKIQYATPEPESYLAIPASSV 954
Db 722 WFEKPLQAYENFKMMHFWGQVDDGDYTNPNSTGH-LHYDNNPNVSSR---PAQLMI 777
Qy 955 DDFDMLTPSSCCRLYISGPNKDKFCPTVNSL--NCL-----KNCMSITMGSVR- 1002
Db 778 LDF-----CQSVRQOP---FVKETLGLMLLPNCPIENLDYMKRCID-DMDSTRK 823
Qy 1003 -----PSVEQFHKYLWFLND-----RPNIKCPK----- 1026
Db 824 DRSPCCDAQFPPEPHI--FEYCLPOSISNMVYDTTFFRGVAGPKFAEAPRLETEDYLGMS 881
Qy 1027 -GGLAAYSTSVNLTS-DGOVLASRE--MAYHKPLKNSODYTEALR-----AAREL 1072
Db 882 GNEAAYSTNGSTFTPLLVKALVIEFESNVAISTYIYANIRQFYSEVHEWFMQOLKTAPPEL 941
Qy 1073 AAN-ITADLRKVPGTDPFAFEVFPYITITNVFYEQYLTLPEGLPMLSLCLVPTFAVSCLLL 1131
Db 942 QCGWFTSDLK-----FYNVQDTLSHDTFVAICLAWAASLAV-LLCF 981
Qy 1132 GLDLRSGLLNLSIVMILVDTVGFMAWLDISYNAVSLINLVSAGMSVFEVSHITRSPAI 1191
Db 982 TVNILISIVAVLTVSISIFNTVAVLILGWQLNILESIAVSTAIGLAVDFSLHYGIHYRM 1041
Qy 1192 STKPTWLERAKEATIS--MGSAPVAGVAMTNLPGILVGLAKAQLIQIFFPRLNLLITLL 1249
Db 1042 SPVKERL-AATQFVLSRIIGPTVMAATT-TGLAGGIMMASNLPIYIQIGVFL--VVVMIV 1097
Qy 1250 GLLHGLVFLPVILSYVGPD-----VNPALALEQKRAEAV 1284
Db 1098 SWFYATFLLMSLLRVAGPQHGLEKWLKWSKSSGSKFYERKPSQVIASEQLLTPSS 1157
Qy 1285 AAWVASCNHPRSVSTADNIYVNHSPFGSIKAGAI SNFLPN 1328
Db 1158 AIVELANSETHELESINSNLIKTIIS---GIESAHALSS-LPRD 1197

```

; GENERAL INFORMATION:
; APPLICANT: Universit. t Zrich
; TITLE OF INVENTION: Dispatched Polypeptides
; FILE REFERENCE: D. melanogaster dispatched sequence
; CURRENT APPLICATION NUMBER: US/10/168,428
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 09/468,237
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1218
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-168-428-2

Query Match 3.1%; Score 212.5; DB 4; Length 1218;
Best Local Similarity 18.4%; Pred. No. 2e-08;
Matches 208; Conservative 184; Mismatches 409; Indels 329; Gaps 48;

QY 374 ELTTPDVELMSPNSQARSEKAFHQHFGPPFRTNQVILTPAPNRSSRYDSLLGPKNFS 433
DB 222 QARTLPTDGMHRRQPIEGFFCD-----SSPRKEYSHFVVQRIGPNATD 267
QY 434 GILDLDLLELELQERLHLQVWSPQARNISLQDIC-----YAPLNPDNTSLYDCC 486
DB 268 SLFDNLGLLAMCOLQDQITEVPSYRAFCEPEMLTTECCRPWSLNPYAAMLANKSCFD-- 325
QY 487 INSLQYFQNNRTLLLTANQTLQMTQSQVDKHF-----LYCANAPLTKDGTALAL 540
DB 326 -----LTTEDVTSLTLLGCEYFHDLMKMDNHCNEIP----- 358
QY 541 SCHADYGAPVFFLAIGGYGKDYSEAEALIMTESLNNYPAGDPRLAQAKLWEEAFLEEM 600
DB 359 HCRAP-----BECKRLNVFVNLNF-----LTDPSFIKSN 388
QY 601 R-----AFORMMAGFQVTFATERSLEINRTTAE DL----- 633
DB 389 DSNVYLKYAMIFIPVAQSNRLPLFHEWEDVE--LINELVEVVAMDGLGLENLFLNELLIT 446
QY 634 PIPATSYIVIFLYISLALGYSWSRWVWVDSKATLGLGVAVVGLVMAAMGFPSVLGIR 693
DB 447 DVMLVSLGDTFVMAWSVLYTGSAFITLMSVCAICFSL-GLAYFFFAIVLVEFFFPYNNLL 505
QY 694 SSIVILQVVPFLVLSGADNIFIVLEY-----QRLPRR-----PGE 730
DB 506 AVVVI-----IGGADVDVFLKIVHCVLTERESNCTLTQSQSALPTLENSDHT 557
QY 731 PREVHIGRALGRVAPSMGLCSLEAICFFLGALTPMPAVRTFALTSGLAVIDLFLQMSA 790
DB 558 SLENIMALTWRHAASMEFTVSLTAGAFVASYSSSITAICFCGIFAGTVVVTNYLLMITW 617
QY 791 FVALLSDSKRQASRLDVCCCKVQBELPPGQEGILLGFFOKA---YAPFLHWHITRG 847
DB 618 LPASVST-MERLFPATMS---CHHPMSIK-----LIHACKKSINRCQMPFECITKS 665
QY 848 VV--LLLFALFGVSLVSMCHISVGLDQELALPKDSYLLDYFLFLNARY-FEVGAPVYFVT 904
DB 666 IMVAYLWLLIFG-ALGASSAVIVFWYVGLQLEKSH---FQLFVSKHPPEV-----YS 715
QY 905 TLGYNFSSEA-----GMNAICSSAGCNPFSTQKIQYATEPPEQSYLAI 948
DB 716 SLKQQFWEKFPQAVENFKMMHMFVWGVAVDGDYTNPNNSYGH-LHYDNNFNVSRR--- 771
QY 949 PASWVDDFDLWLTSSCCRLYISGPNKDKFCSTVNSL--NCL-----KNCMSIT 997
DB 772 PAQLWILD-----COSVROOP-----FYKETLGMLLPNCFIENLIYMKRCID-D 817
QY 998 MGSVR-----PSVEOFHKYLPWFND-----RENKICPK----- 1026
DB 818 MDSTRKDRSPCCDAQFFPEPHI--FEYCLPQISNMWYDTTFFRGVAGPKFAEAPRLETE 875
QY 1027 -----GGLAAYSTSVNLTS-DGQVLASRF---MAYHKPLKNSQDYTEALR----- 1067

; RESULT 91
; US-10-094-749-2651
; Sequence 2651, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTYUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; PRIOR FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2651
; LENGTH: 1203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2651

Query Match 3.0%; Score 210.5; DB 4; Length 1203;
Best Local Similarity 18.5%; Pred. No. 2.9e-08;
Matches 181; Conservative 153; Mismatches 377; Indels 269; Gaps 39;

QY 387 NSQARSEKAFHQHFGPPFRTNQVILTPAPNRSSRYDSLLGPKNFSGLDLDLLELLE 446
DB 6 NSRIRS-----HPO-FGDLQRTTAASCCPSWTGLNIAILNNRSSCQKIVERDVSHLKL 60
QY 447 LOERLRHLQ-----VWSEPAQRNTSLQDICYAPLNPDNTSLYDCCINSLQYFQNNRT 499

```

Db 61 LRTCAHYQNTGLGPDWMAARRKQOLK-----CTNVPKCTKKNV 103
Qy 500 LLLLTANTLMGOTSQVWDKDFLYCAN--APLTFKDGDTALALSCMADYAGPVPFLAIG 557
Db 104 YQIL-----HYLVKDFWTPKTI-----ADYATPALKYSMPLP 134
Qy 558 GYKGDYSEALIMTSIANNYPAGDPRLAQAKLWEAFLEEMRAFORWAGHFOVTTFTA 617
Db 135 SPTEKGS-----MMNLYLDF-----ENWSS-----DGVTTIT-- 164
Qy 618 ERSLEDEINTTAEDLPFATSYIVIFLYISLALGSYSSWSR--VMVDSKATILGLGVAV 675
Db 165 --GLEFGKSLFQDYLMDTVPAIAIVILLVMCVYTKSMFILTMTFAIISLIVSY 222
Qy 676 VLGAVMAGFFSYGLRSLSVLQVVPFLVLSVGADNIFIF--VLEYORLPRRPGPRE 733
Db 223 FLYRVVPHFFPPMNL-TALIIL-----VGIGADDAFVLCVWVNYTKFDPHAETSE 274
Qy 734 VHIGRALGRVAPSMLLCSLEAICFFLGALTPMPAVRTFALTSGLAVIDFLQMSAFVA 793
Db 275 T-VSITLQHAALSFMFTSFTTAAAFYANYVSNITAIRCFGYAGTALVNYLVMVTLPA 333
Qy 794 LLSLDSKROASRLDVCCV-KQBELPPPGQ-----EGLILG-----FPOKAY 836
Db 334 VVVL-----HERYLLNFTCFKKQOQIYDNKSCWTACQKCHKVLPFAISEARIFFEKV 389
Qy 837 APFLHWTIRGVLLFLALFGVLSYMSCHISVGLDQELALPK-----DS 881
Db 390 PCIVIKF--RYLWLFALTVGGAYIVC-----INPKMLPSLELSEFOVFRSSHPFER 442
Qy 882 YLLDY-FLFLNRFEVGAPVYVTTILGYNFSE-----AGNNAICSSA 923
Db 443 YDAEYKGLFMRVHGEELHMPITVINGVSPEDNGNPLNPKSGKGLTLDSSFNIPAS 502
Qy 924 GCNFFSFKIOYATEP--PEQSYLAIPASSWDDFDLWLTSSC-----CRLYISG 973
Db 503 QAWILHFCQKLRNQTFFYQDEQDF-----TSCFIETFKQWENQDCDEPALYPCCSHWSF 558
Qy 974 PNKDFECPSTVNSLNLKNCMSITMGSVRPSVQFHKYLPWFLNDRPNKCPKGGLAAYS 1033
Db 559 FYKQ-----BIFE-----LCIKRAIMELE 577
Qy 1034 TSNVLTSDGVLASRWAYHKLKNSQDYTEALRAA-----REL 1072
Db 578 RSTGYHLDSKTPGPR-----DINDTIRAVVLEFQSTYLETAYEKMHQFYKEV 626
Qy 1073 AMNITADLRKVP-GTDPAPVPPYTTNV-FYEQYLTILPEGLFMLSCLVPTFAVSCLL 1130
Db 627 DSWISSELSAPGLSNG-----WFSNLEFYDLQDSLSDGTILAMGLSVAVAFSV-MLL 680
Qy 1131 LGLDLSGLNLNLSIVMILVDYTFGMALWDISYNAVSLNLVSAVGMSVEFVSHITRSPA 1190
Db 681 TTWNIISLYAIIISAGTIPFTVGSVLLGWELNLVLESVTSIVAVGLSVDFAHYGVAVR 740
Qy 1191 ISTKPTWLERAKEATISM--GSAVPAVAMTNLPGILVGLAKAQLIQIFFPRMLNLT 1247
Db 741 LAPDP--DREGKVIFSLRSGSAMAMAALTTPVAGAMWMPSTVLAAYTOLGTGM--MLIM 795
Qy 1248 LLGLLHGLVPLVILSVGP 1267
Db 796 CISWAPATTFQCMCRCLGP 815

RESULT 92

US-10-311-623-3
; Sequence 3, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda

Query Match 2.9%; Score 202; DB 4; Length 1124;
Best Local Similarity 18.9%; Pred. No. 1.5e-07;
Matches 161; Conservative 134; Mismatches 331; Indels 228; Gaps 35;
Qy 519 KDFLYCANAP-----LTFKDGDTALALSCMADYAGPVPFLAIGGYKGD 563
Db 6 KQO-LACTNVPKCTKKNVYQILHYLVKDF--FMTPKTADYATPALKYSMPLFSPTKEG 61
Qy 564 YSEAEALIMTFSIANNYPAGDPRLAQAKLWEAFLEEMRAFORWAGHFOVTTFTAERSLED 623
Db 62 ES-----MMNLYLDF-----ENWSS-----DGVTTIT--GIEF 89
Qy 624 EINRTTAEDLPFATSYIVIFLYISLALGSYSSWSR--VMVDSKATILGLGVAVLGVAVM 681
Db 90 GIKHSLFQDYLMDTVPAIAIVILLVMCVYTKSMFILTMTFAIISLIVSYFLYRVV 149
Qy 682 AAMGPPSYLGISSVLQVVPFLVLSVGADNIFIF--VLEYORLPRRPGPREVHIGRA 739
Db 150 FHFEPPFPMNL-TALIIL-----VGIGADDAFVLCVWVNYTKFDPHAETSET-VSIT 200
Qy 740 LGRVAPSMLLCSLEAICFFLGALTPMPAVRTFALTSGLAVIDFLQMSAFVALLSLDS 799
Db 201 LQHAALSFMFTSFTTAAAFYANYVSNITAIRCFGYAGTALVNYLVMVTLPAVVL-- 258
Qy 800 KRQEASRLDVCCV-KQBELPPPGQ-----EGLILG-----PFQAYAPFLH 842
Db 259 --HERYLLNFTCFKKQOQIYDNKSCWTACQKCHKVLPFAISEARIFFEKVLPVIX 316
Qy 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPK-----DSYLLDY-- 886
Db 317 P--RYLWLFALTVGGAYIVC-----INPKMLPSLELSEFOVFRSSHPFERDYAEYK 369
Qy 887 FLFLNRYFEVGAPVYVTTILGYNFSE-----AGNNAICSSAGCNF 929
Db 370 KLFMFERVHGEELHMPITVINGVSPEDNGNPLNPKSGKGLTLDSSFNIPASQAWILH 429
Qy 930 FTQKIOYATEP--PEQSYLAIPASSWDDFDLWLTSSC-----CRLYISGPNKDF 979
Db 430 FCQKLRNQTFFYQDEQDF-----TSCFIETFKQWENQDCDEPALYPCCSHWSFPYKQ-- 483
Qy 980 CPTVNSLNLKNCMSITMGSVRPSVQFHKYLPWFLNDRPNKCPKGGLAAYSTSVNLT 1039

Db	484	-----EIFE-----LCIKRINMELERSTGYH 504	
Qy	1040	SDGOVLASREYAHKPLKNSODYTEALRAA-----RELAANITA 1078	
Db	505	LDSTKPGPRF-----DINDTIRAVLEFQSTYLFYLAWEKHOFYKEVDSMISS 553	
Qy	1079	DLRKVP-GTDPAFEPFYITNV-FYEQYLTLPGLFMLSCLVPTPAVSCLLIGLDLR 1136	
Db	554	ELSSAPEGLSNG-----NFVSNLEFYDLQDSLSDGTLIAMGLSVAVAFSV-MLLTWNII 607	
Qy	1137	SGLLNLISVIMLVDTVGFEMALWDISYNAVSLINLVSAVGSVVFVSHITSFALSTKPT 1196	
Db	608	ISLYAIISAGTIFVTGSLVLLGWELNVLESVTISVAVGLSVDPFAVHYGAYRLAPDP- 666	
Qy	1197	WLERAKEATIS--GSAVFAGVAMTNLPGLVLGLAKAQIQLIFFPRLNLIITLLGLH 1253	
Db	667	--DREGKVFISURVGSANMAALITTFVAGAMMFSTVLAYTQLGTFM--MLINCISWAF 722	
Qy	1254	GLVFLFVLISVYGP 1267	
Db	723	ATFFQCMCRCLGP 736	
RESULT 93			
US-10-085-198-168			
; Sequence 168, Application US/10085198			
; Publication No. US2004000907A1			
; GENERAL INFORMATION:			
; APPLICANT: Alsobrook et al.			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same			
; FILE REFERENCE: 21402-279			
; CURRENT APPLICATION NUMBER: US/10/085,198			
; CURRENT FILING DATE: 2002-02-25			
; PRIOR APPLICATION NUMBER: 60/271,646			
; PRIOR FILING DATE: 2001-02-26			
; PRIOR APPLICATION NUMBER: 60/276,401			
; PRIOR FILING DATE: 2001-03-16			
; PRIOR APPLICATION NUMBER: 60/311,981			
; PRIOR FILING DATE: 2001-08-13			
; PRIOR APPLICATION NUMBER: 60/312,858			
; PRIOR FILING DATE: 2001-08-16			
; PRIOR APPLICATION NUMBER: 60/271,840			
; PRIOR FILING DATE: 2001-02-27			
; PRIOR APPLICATION NUMBER: 60/277,324			
; PRIOR FILING DATE: 2001-03-20			
; PRIOR APPLICATION NUMBER: 60/286,096			
; PRIOR FILING DATE: 2001-04-21			
; PRIOR APPLICATION NUMBER: 60/299,695			
; PRIOR FILING DATE: 2001-06-20			
; PRIOR APPLICATION NUMBER: 60/315,614			
; PRIOR FILING DATE: 2001-08-29			
; PRIOR APPLICATION NUMBER: 60/272,405			
; PRIOR FILING DATE: 2001-02-28			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 653			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 168			
; LENGTH: 1561			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-085-198-168			
Query Match 2.9%; Score 197.5; DB 4; Length 1561;			
Best Local Similarity 19.5%; Pred. No. 6e-07;			
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;			
Qy	576	LNNYPAGDPRIAQAUKLWEAFLEEMRAFORMAGMFQVTTAERSLED----- 623	
Db	574	LPNYSVDDR-----WEE-----ORAKFOSFVVTYVAMLAKOSTSKVQVLYGGTD 618	
Qy	624	-----EINRTAED--LPFATSYIVIFLYTISLALGSYSSRWVMDSKATLGLGG-VAVV 676	
Db	619	LFDYEVRTFNMDMLAFISSCIAALVILTSCSVLSFPGI-----ASIGLSCLVLF 673	

RESULT 94
US-10-239-316-17
; Sequence 17, Application US/10239316
; Publication No. US2003012523A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun

Qy	677	IGAVMAANGFSSYLGRSSLVILQVVPFLVLSVCADNIFIVLVEYQRLPR-RPCEPREVH 735	
Db	674	LYHVVFGI---QVLLGILNG-----VAAPFVIVGIGVDVDFVINTYRQATHLEDQLRMH 725	
Qy	736	IGRALGRVAPSMILCSLSEAI CFFLGALT PMPAVRTPALTSGLAVI---LDLFLQMSAFV 792	
Db	726	TVQTAGK---ATFTSTLTAAAYANVPSQIPAVHD FGLFMSLIVSCCWLAVLVTMPAAL 782	
Qy	793	ALLSLSKROEASRLDVC--CCVKPQBLPPPGQ----- 823	
Db	783	GLWSLYLAPLESSQTSCHQNCRSKTSILHPFGDVFATPEQVGGSPAQAPIFYLDDDIPLL 842	
Qy	824	-----GEGLLGFFQKAYAP-----FLLHWI-----TRGVVL 850	
Db	843	EVBEPPVSLGVDVSLVSVSPEGLOPASNTSGRGLHIVLOQLLHHWVAWAKSRWIV 902	
Qy	851	LLFLALFGVSL-----YSM-----CHISVGLDQE- 874	
Db	903	GLFVSIILSVLPASRLRPASRAPLLPRPDNIQVILLDLKYNLSAEGISCITCSGLFOEK 962	
Qy	875	-----LALPK-----DSYLLDY--FLFLNRYFEVGPVFTVLGYNF 910	
Db	963	PHSLQNNIRTSLEKKRSGSGVPMASRPEATLQDPFGTVIISKVKSQGHPA--VYRLSLNA 1020	
Qy	911	SSEAGMNAICSSAG-CNNF-----SFTOKIQVATEFPEQSYLAIPASSWVDDFIDLWL 961	
Db	1021	SLPAPWQAVSPGDGEVPSFQVYRAPFGNFTKLTACMTGVLQLQNASPSRKWM-----L 1074	
Qy	962	TPSSC-----CRLYIS--GP 974	
Db	1075	TTLACDAKRWKDFDFYVATKEQHTKLYFAQSHKPPFHGRVCMVPPGCLSSSPDGP 1134	
Qy	975	NKDFCPSVTNSLNCNKCMSTMG-----SVRPSVEGFHKYLPWF-----LN- 1017	
Db	1135	TKGFE---FVPSEKVPKARLSATFGNFCVNTGCGKPAVRPLVDTGAMVVFVFGIIGNVR 1191	
Qy	1018	-----DRPNIKCPKGLAAYSTSVNLTSD-----GOVL 1045	
Db	1192	TRQVDNHVIGDP--GSVYVDSFDLFKEIGHLCHLCKAIAANSELVKPGCAQLPSGYSI 1249	
Qy	1046	ASRFMAYHKPLKNSQD-----YTEALRAARELAANITADLRKKVGTDPAPFV-P 1094	
Db	1250	SSFQWLHPECKELPEPNLLPQLSHGAVGVRGREGVQVEISMASESTTYKGS-SFQTSYSD 1308	
Qy	1095	YTITNVFEQYLTILPEG-----LFM-----LSLCL--VP 1122	
Db	1309	YLRWESFLQQQLQALPEGSVLRRGFTCEHWKQIFMEIVGVQSALCGLVLSLLICVAVA 1368	
Qy	1123	TFAVSCLLGLDLRSGLNLSIVMLVDTVGFMALWDISYNAVSLINLVSAVGSMEVFV 1182	
Db	1369	VFTTHILL-----LLPVLLSILGIVCLLV--TIMYWSGWM--GAVEAISILVGSSEVDYC 1421	
Qy	1183	SHITRSFATS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI----- 1224	
Db	1422	VHLVEGYLLAGENLPHPQAEADARTQRM--RTLEAVRHVGVAIVSSALTTVIATVPLFFC 1479	
Qy	1225	LVLGLAKAQIQLFFPRLNLLTLGLLHGLFPLVILSVGPDVNPALAEQKRAE--E 1282	
Db	1480	IIAPPAK-----FGKIVALNTGVSILYTLTVSTALLGIMAPS-----SFTRTISFLK 1527	
Qy	1283	AVAAVMVA 1290	
Db	1528	ALGAVLLA 1535	


```
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same And
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-17

Query Match      2.8%; Score 195.5; DB 4; Length 1087;
Best Local Similarity 19.5%; Pred. No. 5.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTTAERSLED----- 623
Db 100 LPNYYSVDDR-----WEE-----QRAKFSFVTVVAMLAQKSTSKVQVLYGGTD 144
QY 624 -----EINRTAED--LPIFATSYIVIFLYISIALGYSYSSWSRVWVDSKATLGLGG-VAVV 676
Db 145 LFDYEVRRFTFNNDMLAFITSSCIAALVYILTSCSVFLSFPGI-----ASIGLSCLVALF 199
QY 677 LGAVMAAMGFPSYLGIRSSVLIVQVPPFLVSLVGADNIFIVLEYQRLPR-RGEPREYH 735
Db 200 LKHVFGI---QYLGILNG-----VAAFVIGVIGVDDVFFVINTYQATHLEDPQRLMTH 251
QY 736 IGRALGRVAPSMLLCSLSEAICFFFLGALTMPAPVARTFALTSLGLAVI---LDFLLOMSAFV 792
Db 252 TWQTAGK---ATFTSLTITAAAYAAVFSQIPAVHDFGLFMSLVSCCVLAVLVTMPAAL 308
QY 793 ALLSDSKRQEARLDVC--CCVKQDELPPGQ----- 823
Db 309 GLWSLYLAPLESSCQTSCHQNSRKTSLHPPGDVFAAPQVGGSPAQGPYPYLDLDDIPLL 368
QY 824 -----GEGLLGLFFOKAYAP-----FLLHWI-----TRGVYL 850
Db 369 EYEEEPVSLLELGDVSLVSVSPGLQPAASNTGSRGHLIVQLQELHHWLVMSAVKSRWIV 428
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
Db 429 GLFVSLIILSLVPSLRPASRAPLLFRPDTHIQVLLDLKYNLSAEGISCITCSGLFQBK 488
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFEVGPVYFVTTILGYNF 910
Db 489 PHSLQNNIRTSLEKKRRGSGVPWASRPEATLQDFPGTVVISKVKSGHPA--VYRLSLNA 546
QY 911 SSEAGHNAICSSAG-CNPF-----STQKIQYATEBPQSYLAI PASSWDDPIDWL 961
Db 547 SLPAWQAVSPGDGEVPSPQYRAPFGNFTKXLTACMSTVGLLQASPSRKNW-----L 600
QY 962 TPSSC-----CRLYIS--GP 974
Db 601 TTLACDAKKGWKFDSFYVATKEQHQTRKLYFAQSHKPPFHGRVCMAPPGCCLLSSPDGP 660
QY 975 NKDKFCFSTVNSLNCIKNCSITMG-----SVRPSVEQPHKYLPMF-----LN- 1017
Db 661 TKGFF---FVPSEKVPKARLSATFGFNPVNTGCGKPAVRPLVDITGAMVFFVFGIIGVNR 717
QY 1018 -----DRPNIKCPKGGLAAYSTSVNLTSD-----GQVL 1045
Db 718 TRQVDNHVIGDP--GSVYDSSFDLFKEIGHLCHLCKAIAANSSELVKGPGAQCLPSGYSI 775
QY 1046 ASRFMAYHKPLKNSQD-----YTEALRAARELANITADLRKVPGTDPAPFVP-P 1094
Db 776 SSFLQMLHPECKELPEPNLLPQOLSHGAVGVRGVQWISMAFESTYTKGS-SFQTSYSD 834
QY 1095 YTITNVFYEQYLITLPEG-----LFW-----LSLCL--VP 1122
```

```
Db 835 YLRWESFLQOQLOALPEGSVLRGRFQTCBHWKQIFMEIVGVQSALCGLVLSLICVAANA 894
QY 1123 TPAVSCLLGLDLRSGLNLLSTVMILVDTVGFMALWDISYNAVSLINLVSAYGMSVFEV 1182
Db 895 VFTTHILL-----LLPVLISILGIVCLVW-TIMYWSGEM--GAVEAISILVGSVDYC 947
QY 1183 SHITRSPAIS-----TKPTWLERAKGATISMGSAVFAGVAMTNLPGI----- 1224
Db 948 VHLVEGYLLAGENLPHQAEADARTQW--RTLEAVRHVGVAIVSSALITVTIATVPLFFC 1005
QY 1225 LVVLGAKAQLIOIPFFRLNLLITLGLLHGLPLPVLISYVGVDPNPALALEKQRAE--E 1282
Db 1006 IIAFPK-----FGKIVALTNGVSILYTLTSTALLGIMAPS-----SFTRTTSFLK 1053
QY 1283 AVAAVMVA 1290
Db 1054 ALGNVLLA 1061

RESULT 95
US-10-239-316-40
; Sequence 40, Application US/10239316
; Publication No. US20030125253A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: No. US20030125253A1el Protein, Process for Producing The Same An
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 40
; LENGTH: 1392
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-40

Query Match      2.8%; Score 195.5; DB 4; Length 1392;
Best Local Similarity 19.5%; Pred. No. 7.5e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;

QY 576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTTAERSLED----- 623
Db 405 LPNYYSVDDR-----WEE-----QRAKFSFVTVVAMLAQKSTSKVQVLYGGTD 449
QY 624 -----EINRTAED--LPIFATSYIVIFLYISIALGYSYSSWSRVWVDSKATLGLGG-VAVV 676
Db 450 LFDYEVRRFTFNNDMLAFITSSCIAALVYILTSCSVFLSFPGI-----ASIGLSCLVALF 504
QY 677 LGAVMAAMGFPSYLGIRSSVLIVQVPPFLVSLVGADNIFIVLEYQRLPR-RGEPREYH 735
Db 505 LKHVFGI---QYLGILNG-----VAAFVIGVIGVDDVFFVINTYQATHLEDPQRLMTH 556
QY 736 IGRALGRVAPSMLLCSLSEAICFFFLGALTMPAPVARTFALTSLGLAVI---LDFLLOMSAFV 792
Db 557 TWQTAGK---ATFTSLTITAAAYAAVFSQIPAVHDFGLFMSLVSCCVLAVLVTMPAAL 613
QY 793 ALLSDSKRQEARLDVC--CCVKQDELPPGQ----- 823
Db 614 GLWSLYLAPLESSCQTSCHQNSRKTSLHPPGDVFAAPQVGGSPAQGPYPYLDLDDIPLL 673
QY 824 -----GEGLLGLFFOKAYAP-----FLLHWI-----TRGVYL 850
Db 674 EYEEEPVSLLELGDVSLVSVSPGLQPAASNTGSRGHLIVQLQELHHWLVMSAVKSRWIV 733
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
```

Db 734 GLFVSILSLVFAASRLAPASRAPLLFRPDNTNIQVLLDLKYNLSAEGISCITCSGLFOEK 793
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFVGVAPVYFVTLGLGNF 910
Db 794 PHSLQNNRTSLKKRGSGVWPASRPENTLQDFPGTVYISKVKSQGHPA--VYRLSLNA 851
QY 911 SSEAGNAICSSAG--CNPF-----SFTQKIQVATEFPESQYLAIPASSWDDFDWL 961
Db 852 SLPAQWAVSPGDGVPSPFQVYRAPFGNFTKLTACMSTVGLLQAASPSRKWM-----L 905
QY 962 TPSSC-----CRLYIS--GP 974
Db 906 TTLACDAKRWKFDPSFYVATKEOQHTKLYPAQSHKPPFHGRVCMAPPGCLSSSPGP 965
QY 975 NKDKFCPTSVNSLNCNKQMSITMG-----SVRPSVEQFHLYLPMF-----LN- 1017
Db 966 TKGFF--FVPSKVPKARLSATFCGPNVNTGCGKPAVRPLVDGAMVFFVGIIGNR 1022
QY 1018 ----DRPNIKCPKGLAAYSTSVNLTS-----QVQL 1045
Db 1023 TRQVDNHVIGDP--GSVVYDSSFDFLKEIGHLCHLCKAIAANSSELVKPGGAQCLPSGYSI 1080
QY 1046 ASRFWAYHKPLKNSQD-----YTEALRAARELAANITADLRKVPGTDPAPFVP-P 1094
Db 1081 SSFLOMLHPECKELPEPILLPQLSHGAVGVRGVQWISMAFESTTYKGS-SFQTYSD 1139
QY 1095 YTITNVFYQYLTILPEG-----LFM-----LSLCL--VP 1122
Db 1140 YLRWESFLOQQLQALPEGSVLRGFGTCHEMKQIFMEIVGVQSALCGLVLSLICVAVA 1199
QY 1123 TFVASCLLGLDLRSLNLLSIWMLVDTVGFMAIDSYNAVSLINLVSAVGHVSVEV 1182
Db 1200 VFTTHILL-----LLPVLISILGIVCLV--TIMYWSGWM--GAVEAISLSILVGSVDYC 1252
QY 1183 SHITRSFAIS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI----- 1224
Db 1253 VHLVEGYLLAGENLPPHQEDARTQW--RTLEAVRHVGVAIVSSALTVIATVPLFPFC 1310
QY 1225 LVGLAKAQIQLIFFPRLNLLITLLGLHLGLVPLPVILSVYGVDPVNPALALEKQRAE--B 1282
Db 1311 ITAPPAK-----FGKIVALTNGVSILYTLTVSTALLGIMAPS-----SFTRTTSFLK 1358
QY 1283 AVAANMVA 1290
Db 1359 ALGAVLLA 1366

RESULT 96
US-10-437-963-102679
; Sequence 102679, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, Thomas J.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 102679
; LENGTH: 100
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_100179C.1.pgp
US-10-437-963-102679

Query Match 2.8%; Score 190; DB 4; Length 100;
Best Local Similarity 41.2%; Pred. No. 3.7e-08;
Matches 33; Conservative 24; Mismatches 23; Indels 0; Gaps 0;
QY 1200 RAKEATTSMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLLGLHLGLVPLP 1259
Db 15 RARQALSTMGASVSGITLTKLVGVIVLRFPAKSEVFFVYVQMYLALVIGFLHGLIFLP 74
QY 1260 VILSVYGVDPVNPALALEQR 1279
Db 75 VVLSLCGPPSKVMKPLEQSQ 94
RESULT 97
US-10-168-428-3
; Sequence 3, Application US/10168428
; Publication No. US20030165897A1
; GENERAL INFORMATION:
; APPLICANT: Universit,t Zrich
; TITLE OF INVENTION: Dispatched Polypeptides
; FILE REFERENCE: D. melanogaster dispatched sequence
; CURRENT APPLICATION NUMBER: US/10/168,428
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 09/468,237
; PRIOR FILING DATE: 1999-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 3
; LENGTH: 931
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-168-428-3

Query Match 2.7%; Score 187.5; DB 4; Length 931;
Best Local Similarity 20.2%; Pred. No. 2e-06;
Matches 152; Conservative 119; Mismatches 252; Indels 231; Gaps 36;
QY 606 RMAG-----MPQVTTFAERSLEDEINRTTAEDLPFPATSYI-----VIFLYISLALGS 653
Db 299 RMAGIALANMKNKVFQ-----ERIQDLSLFAAFSA--LLVFSCFLIYSRSIIFTCCIILMVT 352
QY 654 YSSWSRVWVDSKATLGLGVAVVLGAVMAAMGFPSVLGRSSSLVILQVVPFLVLSVGADN 713
Db 353 LSS-----GVAFITVVLGIDFFPFINL-----LVVILISIGADD 389
QY 714 IFIFVLEYQRLPRPFCEPREVHIGR-----ALGRVAPSMMLLCSLEA 755
Db 390 AFLLLVYRREVERMSH--LEYKVGSIYIPLYRESOLLSRSLRSLHSHLSVSMFVTSLTAA 448
QY 756 ICFFLGLTTPMPAVRTFALTSLGLAVLIDFLLOMSAFVALLSLDSK--RQASRLDVCCKVK 814
Db 449 STFLTNLSSPVLRLCFGVYAALTVTNVILVLLPGAILLSRPIRKKLSRGD-----EE 504
QY 815 POELPPPGQEGLLGFFOKAYAPF-----LLHWITRGVLLLLFLALFVSLSYMSCHIS 868
Db 505 PEKI-----ESHYFASKITETHYERFG-IFICSLIMTGLSLFIIFQ-N 547
QY 869 VGLDQELALPK-----DSYLLDYF-----LFNRYFEVGAAPVFTVLGY 908
Db 548 PGLKTPQTNPTKLLVDSNTHIEYFDNNVHFNFQWQSRARLVKNFVFGVDAIKETSLSPY 607
QY 909 NPSSEAGMAICSSACCNPS-----FTQKIQY---ATEPPEQSILAIPAS--SWVDDF 957
Db 608 NKFSK-----NFSQAHYSLDTKLDFYRRIVNLESKKYOLVNYTHVSWADKI 654
QY 958 IDWLTPSSC-----CRLYISGPNKDKFCPTSVNSLNCNKQMSITMGVSRPSVEQ 1007
Db 655 LQ--ANESCFSENKTIHECILSASVRN-----NLHQFPDDFSVIGD-GPIDQ 703
QY 1008 FHKYLPWFLNDRNPNIKCPKGLAAYSTSVNLTSVGQVLASRFMAYH---KPLKNSQDYTE 1064
Db 704 DLKVGYFISIPSNQK-----LQVDTEMIGSFQFQIEESCKQIKNATSDSV 749

Tue Apr 11 10:57:57 2006

us-10-736-769-4.top100.rapbm

```
QY 1187 RSPAISTKP---TWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFFPRLN 1243
Db 966 GNYCTDKKPAINWIE-----GRGKSV---VTEAVIPGAIVEKVLKTTVAAL--VELN 1013
QY 1244 LLITLLGL-----LHGLVFLPVILSVGPDVNPALALEOKRAEEAAVAV 1287
Db 1014 ISKNLIGSMAGSVGGFNAHAANILTAIYLATGQD--PAQNVSSNCITLMKAV 1065
```

```
RESULT 100
US-10-060-756A-14
; Sequence 14, Application US/10060756A
; Publication No. US20030046717A1
; GENERAL INFORMATION:
; APPLICANT: Zhang, Jian
; TITLE OF INVENTION: HUMAN TESTIS EXPRESSED PATCHED LIKE PROTEIN
; FILE REFERENCE: PB0177
; CURRENT APPLICATION NUMBER: US/10/060,756A
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/327,898
; PRIOR FILING DATE: 2001-10-09
; NUMBER OF SEQ ID NOS: 4804
; SOFTWARE: Acomica Sequence Listing Engine
; SEQ ID NO 14
; TYPE: PRT
; LENGTH: 101
; ORGANISM: Homo sapiens
US-10-060-756A-14
```

```
Query Match 2.5%; Score 170; DB 4; Length 101;
Best Local Similarity 36.0%; Pred. No. 2.1e-06;
Matches 36; Conservative 26; Mismatches 36; Indels 2; Gaps 2;

QY 1167 SLINLYSAVGMSEVFVSHITRSPAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILV 1226
Db 1 SMINLVICTGFDFDSAHISYAFVSSSQPSVQKSVKVEALYLLGVPVLQS-AISTIIGVCV 59

QY 1227 LGLAKAQLIQIFFPRLNLLITLLGLHGLVFLPVILSVYG 1266
Db 60 LAAAKAYIFRT-FFKIMFLVMIFGAHGLIFIPVLTFFG 98
```

Search completed: April 11, 2006, 01:01:15
Job time : 232 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:58:02 ; Search time 27 Seconds
(without alignments)
1538.804 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLRLAQ.....GSIKGAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 184161 seqs, 31191982 residues

Total number of hits satisfying chosen parameters: 184161

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Published Applications_AA_New:*
1: /SIDSS/ptodata/2/pubppaa/US08_NEW_PUB.pep:*
2: /SIDSS/ptodata/2/pubppaa/US06_NEW_PUB.pep:*
3: /SIDSS/ptodata/2/pubppaa/US07_NEW_PUB.pep:*
4: /SIDSS/ptodata/2/pubppaa/PCT_NEW_PUB.pep:*
5: /SIDSS/ptodata/2/pubppaa/US09_NEW_PUB.pep:*
6: /SIDSS/ptodata/2/pubppaa/US10_NEW_PUB.pep:*
7: /SIDSS/ptodata/2/pubppaa/US11_NEW_PUB.pep:*
8: /SIDSS/ptodata/2/pubppaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6896	99.8	1332	7	US-11-242-459-9
2	2402.5	34.8	1278	6	US-10-995-561-952
3	942	13.6	194	7	US-11-242-459-8
4	195.5	2.8	1087	7	US-11-242-459-17
5	195.5	2.8	1392	7	US-11-242-459-40
6	191	2.8	127	7	US-11-096-568A-6778
7	173	2.5	85	7	US-11-242-459-35
8	143	2.1	445	7	US-11-242-459-34
9	143	2.1	456	6	US-10-392-234A-18
10	141.5	2.0	1048	6	US-10-392-234A-16
11	140.5	2.0	1046	6	US-10-467-657-2648
12	134	1.9	1067	6	US-10-392-234A-14
13	127	1.8	1048	6	US-10-392-234A-16
14	126.5	1.8	200	7	US-11-242-459-16
15	126	1.8	801	6	US-10-793-626-2020
16	121.5	1.8	1043	6	US-10-392-234A-34
17	120.5	1.7	2426	7	US-11-203-806A-11
18	119	1.7	1048	6	US-10-392-234A-20
19	117.5	1.7	1049	6	US-10-392-234A-12
20	115.5	1.7	914	7	US-11-072-512-2923
21	113	1.6	1034	6	US-10-392-234A-30
22	112.5	1.6	697	7	US-10-392-389-362
23	109.5	1.6	758	7	US-11-096-568A-30412
24	109.5	1.6	772	7	US-11-096-568A-30411
25	109.5	1.6	827	7	US-11-096-568A-30410

1049	1.6	1049	7	US-11-098-686-11239	Sequence 11239, A
1032	1.6	1032	6	US-10-392-234A-67	Sequence 67, Appl
648	1.6	648	7	US-11-087-099-4135	Sequence 4135, Ap
443	1.5	443	7	US-11-087-099-11457	Sequence 11457, A
783	1.5	783	7	US-11-082-389-354	Sequence 354, App
324	1.5	324	7	US-11-087-099-11792	Sequence 11792, A
1080	1.5	1080	7	US-11-096-568A-27723	Sequence 27723, A
1097	1.5	1097	7	US-11-096-568A-27722	Sequence 27722, A
1128	1.5	1128	7	US-11-096-568A-27721	Sequence 27721, A
315	1.5	315	7	US-11-087-099-9820	Sequence 9820, Ap
471	1.5	471	7	US-11-087-099-4362	Sequence 4362, Ap
757	1.5	757	7	US-11-096-568A-27509	Sequence 27509, A
771	1.5	771	7	US-11-096-568A-27508	Sequence 27508, A
826	1.5	826	7	US-11-096-568A-27507	Sequence 27507, A
251	1.5	251	7	US-11-087-099-1098	Sequence 1098, Ap
331	1.5	331	7	US-11-087-099-8831	Sequence 8831, Ap
2042	1.5	2042	6	US-10-506-454-134	Sequence 134, App
2312	1.5	2312	7	US-11-126-313-34	Sequence 34, Appl
463	1.5	463	7	US-11-087-099-3980	Sequence 3980, Ap
475	1.5	475	7	US-11-087-099-2723	Sequence 2723, Ap
324	1.5	324	7	US-11-087-099-5717	Sequence 5717, Ap
647	1.5	647	7	US-11-087-099-9244	Sequence 9244, Ap
726	1.5	726	7	US-11-096-568A-2446	Sequence 2446, Ap
727	1.5	727	7	US-11-096-568A-2445	Sequence 2445, Ap
836	1.5	836	7	US-11-096-568A-2444	Sequence 2444, Ap
279	1.4	279	7	US-11-087-099-4795	Sequence 4795, Ap
389	1.4	389	6	US-10-204-252-10	Sequence 10, Appl
391	1.4	391	6	US-10-204-252-6	Sequence 6, Appl
391	1.4	391	6	US-10-204-252-8	Sequence 8, Appl
391	1.4	391	6	US-10-204-252-12	Sequence 12, Appl
391	1.4	391	6	US-10-204-252-16	Sequence 16, Appl
391	1.4	391	6	US-10-204-252-28	Sequence 28, Appl
577	1.4	577	6	US-10-718-264-3	Sequence 3, Appl
716	1.4	716	6	US-10-506-454-1056	Sequence 3, Appl
391	1.4	391	6	US-10-204-252-14	Sequence 14, Appl
3402	1.4	3402	6	US-10-204-252-18	Sequence 18, Appl
324	1.4	324	7	US-11-087-099-4100	Sequence 4100, Ap
336	1.4	336	7	US-11-087-099-7384	Sequence 7384, Ap
465	1.4	465	6	US-10-793-626-2928	Sequence 2928, Ap
836	1.4	836	7	US-11-087-099-2386	Sequence 2386, Ap
96	1.4	96	6	US-10-858-730-118	Sequence 118, App
479	1.4	479	7	US-11-087-099-4490	Sequence 4490, Ap
479	1.4	479	7	US-11-087-099-10324	Sequence 10324, A
479	1.4	479	7	US-11-087-099-11713	Sequence 11713, A
983	1.4	983	6	US-10-511-538-91	Sequence 91, Appl
1167	1.4	1167	7	US-11-097-125-2	Sequence 2, Appl
338	1.4	338	7	US-11-087-099-8709	Sequence 8709, Ap
686	1.4	686	6	US-10-821-234-1027	Sequence 1027, Ap
2186	1.4	2186	7	US-11-096-568A-28283	Sequence 28283, A
2199	1.4	2199	7	US-11-096-568A-28282	Sequence 28282, A
2301	1.4	2301	7	US-11-096-568A-28281	Sequence 28281, A
325	1.4	325	7	US-11-087-099-9929	Sequence 9929, Ap
461	1.4	461	7	US-11-183-205-32	Sequence 32, Appl
461	1.4	461	7	US-11-260-192-2	Sequence 2, Appl
469	1.4	469	7	US-11-087-099-11321	Sequence 11321, A
552	1.4	552	7	US-11-087-099-10983	Sequence 10983, A
735	1.4	735	6	US-10-505-263-8	Sequence 8, Appl
758	1.4	758	6	US-10-505-263-6	Sequence 6, Appl
793	1.4	793	7	US-11-087-099-2968	Sequence 2968, Ap
231	1.4	231	7	US-11-087-099-4589	Sequence 4589, Ap
467	1.4	467	7	US-11-087-099-2054	Sequence 2054, Ap
477	1.4	477	6	US-10-479-873-9	Sequence 9, Appl
599	1.4	599	7	US-11-087-099-5936	Sequence 5936, Ap
336	1.4	336	6	US-11-205-109-4	Sequence 4, Appl
470	1.4	470	6	US-10-511-538-83	Sequence 83, Appl
1431	1.4	1431	7	US-11-128-059-2	Sequence 2, Appl
296	1.4	296	7	US-11-082-389-16	Sequence 16, Appl
461	1.4	461	7	US-11-087-099-2325	Sequence 2325, Ap
461	1.4	461	7	US-11-087-099-4376	Sequence 4376, Ap
461	1.4	461	7	US-11-087-099-9669	Sequence 9669, Ap
602	1.4	602	7	US-11-087-099-10340	Sequence 10340, A
679	1.4	679	6	US-10-506-454-279	Sequence 279, App

99	93.5	1.4	824	7	US-11-090-617-657	Sequence 657, App
100	93.5	1.4	4868	7	US-11-044-111-24	Sequence 24, App1
ALIGNMENTS						
RESULT 1						
US-11-242-459-9						
; Sequence 9, Application US/11242459						
; Publication No. US20060035835A1						
; GENERAL INFORMATION:						
; APPLICANT: TANIYAMA, Yoshio						
; APPLICANT: KITA, Shunbun						
; APPLICANT: SATOMI, Tomoko Komiya						
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof						
; FILE REFERENCE: 2703USOP						
; CURRENT APPLICATION NUMBER: US/11/242,459						
; CURRENT FILING DATE: 2005-10-03						
; PRIOR APPLICATION NUMBER: US/10/239,316						
; PRIOR FILING DATE: 2002-09-19						
; PRIOR APPLICATION NUMBER: PCT/JP01/02279						
; PRIOR FILING DATE: 2001-03-22						
; PRIOR APPLICATION NUMBER: JP2000-088595						
; PRIOR FILING DATE: 2000-03-24						
; NUMBER OF SEQ ID NOS: 59						
; SEQ ID NO 9						
; LENGTH: 1332						
; TYPE: PRT						
; ORGANISM: Human						
US-11-242-459-9						
Query Match						
Best Local Similarity 99.8%; Score 6896; DB 7; Length 1332;						
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;						
QY	1	MAEAGLRGMLLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN	60			
DB	1	MAEAGLRGMLLWALLRLAQSEPYTHIQPGYCAFYDECGKPELSGSLMTLSNVCSLSN	60			
QY	61	TPARKITGDHLLILLOKICRPLTYGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF	120			
DB	61	TPARKITGDHLLILLOKICRPLTYGPNTOACCSAKQLVLEASLSITKALLTRCPACSDNF	120			
QY	121	VNLHCHNTCSPNQSLFINTVRAQLGAGQLPAVVAYEAFYQHSFABQSYDSCSRVRPAA	180			
DB	121	VNLHCHNTCSPNQSLFINTVRAQLGAGQLPAVVAYEAFYQHSFABQSYDSCSRVRPAA	180			
QY	181	ATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITPHLEPGQAVGSGIQLNEGV	240			
DB	181	ATLAVGTMCGVYGSALCNAQRWLNFGQDTGNGLAPLDITPHLEPGQAVGSGIQLNEGV	240			
QY	241	ARCNESGGDDVATCSQDCQCAASCPAIARPOALDSTFVLGQMPGSLVLIILCSVFVAVTI	300			
DB	241	ARCNESGGDDVATCSQDCQCAASCPAIARPOALDSTFVLGQMPGSLVLIILCSVFVAVTI	300			
QY	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV	360			
DB	301	LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVLSV	360			
QY	361	IPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420			
DB	361	IPVVALAAGLVFTLTDTPVELMSAPNSQARSEKAFHDQHPGPFRTNQVILTAPNRSY	420			
QY	421	RYDSLILGPKNFGSILDLLELELQERLRHLQVNSPEAQRNISLQDICYAPLNPNT	480			
DB	421	RYDSLILGPKNFGSILDLLELELQERLRHLQVNSPEAQRNISLQDICYAPLNPNT	480			
QY	481	SLYDCCINSLLIQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGITAL	540			
DB	481	SLYDCCINSLLIQYFQNNRTLLLTANQTLMGQTSQVDWKQHFLYCANAPLTFKDGITAL	540			
QY	541	SCMADYGAPVFPFLAIGYKGDYSEAEALIMTSLNNYPAGDPRLAQAKLWEEAFLEEM	600			

1046 ASRFMAVHKPLKNSQD-----YTEALRAARELAANITADLRKVPCTDPAFEVFP-P 1094
776 SSFLOMLHPECKELPEPNLLPGQLSHGAVGRVQWISMAFESTTYKGS-SFQTYSD 834
1095 YTTITNPFYEQYLITLPEG-----LFM-----LSLCL--VP 1122
835 YLRWESFLOQOLALPEGSVLRGRFQTCHEHWKQIFMBEIVGVQSGALCVLSLITCVAAVA 894
1123 TFAVSCLLGLDLRSGLNLSLIMVILDTVGMALWDISYNAVSLINLYSAVGMSVEFV 1182
895 VFTHHILL-----LLPVLLSILGIVCLVV-TIMVSGWEM--GAVEALSLSLTVSSVDYC 947
1183 SHITRSEAFIS-----TKPTWLBRAKEATISMSGSAVFAGVAMTNLPGI----- 1224
948 VHLVEGYLLAGENLPPHQAEARTQRM--RTLEAVRHVGVAIVSSALTTVIATVPLPFC 1005
1225 LVLGLAKAQLIQIFFRNLNLLITLLGLHLGVFLPVLISYVGPDPNPALEQKRAE--E 1282
1006 IIAFFAK-----FGKIVALNTGVSILYTLTVSTALLGINAPS-----SPTRTRTSFLK 1053
1283 AVAAMVVA 1290
1054 ALGAVLLA 1061
RESULT 5
US-11-242-459-40
; Sequence 40, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITAH, Shunbun
; FILE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; CURRENT APPLICATION NUMBER: 2703USOP
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 17
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-17

Query Match 2.8%; Score 195.5; DB 7; Length 1087;
Best Local Similarity 19.5%; Pred. No. 6.1e-07;
Matches 200; Conservative 129; Mismatches 320; Indels 379; Gaps 47;
QY 576 LNNYPAGDRLAQAKLWEEAFLEEMRAFORRMAGMFQVTFRTAERSLED----- 623
DB 100 LPNYYSVDDR-----WEB-----ORAFQSFVTVYVAMLAQKSTSKVQVLYGGTD 144
QY 624 ----EINRTTAED--LPFATSYVIVFLYISALGSSYSSWSRVMVDSKATLGLGG-VAVV 676
DB 145 LFDYEVRRFTNNDMLLAFISSCIAALVILTSVFLSFFGI-----ASIGLSCLVALP 199
QY 677 LGAVMAAMGFFSYLGRSSVLQVVPFLVSVGADNIFIVLEVYQRLPR-RPGEPREVH 735
DB 200 LYHVVFGEI---QYLGILNG-----VAAFVIVGIGVDDVFVINTYRQATHLEDQPLRMH 251
QY 736 IGRALGRVAPSMMLCSISEAICFFLGAITPMPAVRTFALTSGLAVI---LDFLQMSAFV 792
DB 252 TVTAGK---ATFTSITLTTAAVAAVNFVSIQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308
QY 793 ALLSLDSKQBPASRLDVC--CCVKPQELPPPGQ----- 823
DB 309 GLWSLYLAPLESQTSCHQNCRSKTSLHFFGDVFAAPEQVGGSPAQGPPIYLDLDDIPL 368
QY 824 -----GEGLLGFTQKAVP-----FLHVI-----TRGVVL 850
DB 369 EVEBEPVSLGLGVSLVSVSPGELQPASNTGSRGHLIVQLQELHWHVLSAVKSRVIV 428
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDQE- 874
DB 429 GLFVSILILSLVFASRLRPSRAPLLRPDNTIOVLDLKYNLSAEGISCTCSGLFQEK 488
QY 875 -----LALPK-----DSVLLDY--FLFLNRYEVEGAPVYVTTLYGNF 910
DB 489 PHSIQNIRTSLEKRRGSGVPMASRPENTLQDPGTVYSKVKSQGHPA--VYRLSLNA 546
QY 911 SSEAGMNAICSAG--CNNF-----SFTQKIQYATEFPFQSYLAIPASSWDDDFDML 961
DB 547 SLPAWQAVSPGDGEVPSFQVYRAPFGNFTKCLTACMTSVGLLQAAQSPSRKWM-----L 600
QY 962 TPSSC-----CRLYIS--GP 974
DB 601 TTLACDAKRWKPFDSFVATKEQHTKLYPAQSHKPPFHGRVCMAPPGLLSSPDGP 660
QY 975 NKDKFCPSTVNSLNLCKNCSITMG-----SVRPSVEQFHKLWPF-----LN- 1017
DB 661 TKGFF---FVPSKYPKARLSATFGFNCVNTGCGKPAVRPLVDGTGMVVFVFGIGVNR 717
QY 1018 ----DRPNKCPKGGGLAAYSTSVNLTSD-----GQVL 1045
DB 718 TRQVDNHWIGDP--GSVVYDSSFDLFXEIGHLCHLCKAIAANSELVKPGGGAQCLPSGYSI 775

QY 824 -----GEGLLGPPQKAYAP-----PILHWI-----TRGVVL 850
Db 674 EVVEEFVSLGDSVLSVSPGLOPASNTGSRGHLIVQLQELHHWVLSAVKSRWIV 733
QY 851 LLFLALFGVSL-----YSM-----CHISVGLDGE- 874
Db 734 GLFVSLILSLVSPASRLRPASRAPLLFRPDNTIQVLLDLKYNLSAGISCTCSGLFQEK 793
QY 875 -----LALPK-----DSYLLDY--FLFLNRYFEVGPVYFVTTLGYNF 910
Db 794 PHSLQNNIRTSLEKKRSGSVWASREATLQDPFGTVVISKVSQHPA--VYRLSLNA 851
QY 911 SBEAGNNAICSSAG--CNWF-----SPTQIQYATEPPEOSYLAIPASSWDDIDWL 961
Db 852 SLPAPQAVSPGDGPEVSPQVYRAPFGNETKKLTACMSTVGLIQAASPSRKWM-----L 905
QY 962 TPSSC-----CRLYIS--CP 974
Db 906 TTLACDAKRGKWFDFSYVATKEQQHTRKLYFAQSHKPPFHGRVCMWAPGCLLSSPDGP 965
QY 975 NKDKFCPTVNSLNCNKMSITMG-----SVRPSVEQPHKYLPMF-----LN- 1017
Db 966 TKGFP---FVPSKVPKARLSATFGFNPVNTCCGKPAVRPLVDTGAMVVFVGIIGVNR 1022
QY 1018 ---DRPNIKCPKGLAAYSTVNLTSD-----GQVL 1045
Db 1023 TRQVDNHVIGDP--GSWYDSSFDLPKEIGHLCHLCKATAANSSELVKPGGAQCLPSGYSI 1080
QY 1046 ASRFMAYHKLNSQD-----YTEALRAARELANITADLRKVPCTDPPEVP-P 1094
Db 1081 SSFLQMLHPECKELPFPNLLPQOLSHGAVGVREGRVQWISMAFESTYTKGKS--SPQTSYSD 1139
QY 1095 YRITNVFYQYITLPEG-----LFM-----LSLCL--VP 1122
Db 1140 YLWESFLQQLQALPEGSVLRGPTCEHWQIFMEIVGVQSALCGLVSLICVAVA 1199
QY 1123 TPVASCLLGLDLRGLNLLSIYMLVDTVGFMALWDISYNAVSLINIVSAGMSVERV 1182
Db 1200 VETTHILL---LLPVLISILGIVCLV--TIMWSGWEM--GAVERAISLSILVGSVDYC 1252
QY 1183 SHITSPAIS-----TKPTWLERAKEATISMGSAVFAGVAMTNLPGI-----1224
Db 1253 VHLVGEYLLAGENLPPHQAEDARTORQW--RTLEAVRHVGVAIVSSALTVTIATVPLFPFC 1310
QY 1225 LVVLGLAKAQLIQIFPFRNLTLTLLGLHGLVFLPVLSYVGDVNPALALEQKRAE--E 1282
Db 1311 ITAPFAK-----FGKIVALMTGVSILYTLTVSTALLGIMAPS-----SFRTRTSFLK 1358
QY 1283 AVAAVMVA 1290
Db 1359 ALGAVILA 1366

RESULT 6
US-11-096-568A-6778
; Sequence 6778, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; PRIOR FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6778
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(127)

; OTHER INFORMATION: Ceres Seq. ID no. 15168324
US-11-096-568A-6778

Query Match 2.8%; Score 191; DB 7; Length 127;
Best Local Similarity 44.3%; Pred. No. 6.4e-08;
Matches 35; Conservative 21; Mismatches 23; Indels 0; Gaps 0;
QY 1199 ERAKEATISMGSAVFAGVAMTNLPGILVGLAKAQLIQIFPFRNLTLTLLGLHGLVFL 1258
Db 34 QRAKTALCTMGASVFSGITLTKLVGLVLCFSTSQIFVYVYFQMYLALVLIGFLHGLVFL 93
QY 1259 PVLSYVGDVNPALALEQ 1277
Db 94 PVVLSLFGPPRYTVIKEQ 112

RESULT 7
US-11-096-568A-6779
; Sequence 6779, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 6779
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(85)
; OTHER INFORMATION: Ceres Seq. ID no. 15168325
US-11-096-568A-6779

Query Match 2.5%; Score 173; DB 7; Length 85;
Best Local Similarity 44.3%; Pred. No. 9.6e-07;
Matches 31; Conservative 19; Mismatches 20; Indels 0; Gaps 0;
QY 1208 MGSAVFAGVAMTNLPGILVGLAKAQLIQIFPFRNLTLTLLGLHGLVFLPVLSYVGP 1267
Db 1 MGASVFSGITLTKLVGLVLCFSTSQIFVYVYFQMYLALVLIGFLHGLVFLPVLSLFGP 60
QY 1268 DVNPALALEQ 1277
Db 61 PLRYTVIKEQ 70

RESULT 8
US-11-242-459-35
; Sequence 35, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 35
; LENGTH: 445
; TYPE: PRT

793 ALLSLDSKQEQASRLDVC--CCVKQELPPPG 822
309 GLWSLYLAPLESSCQTSCHQNCRSKRTSLHFP 340

RESULT 10
US-10-392-234A-18
; Sequence 18, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Kiazni, Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 1048
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
US-10-392-234A-18

793 ALLSLDSKQEQASRLDVC--CCVKQELPPPG 822
309 GLWSLYLAPLESSCQTSCHQNCRSKRTSLHFP 340

Query Match 2.1%; Score 143; DB 7; Length 445;
Best Local Similarity 25.4%; Pred. No. 0.0024;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVFTFAERSLED----- 623
100 LPNYSVDDR-----WEE-----QRAKFSFVTVVAMLAQKSTSKVQVLYGGTD 144

624 ----EINRTTAD--LPFATSYIVIFYISLALGSYSSWRVMDSKATLGLGG-VAVV 676
145 LPDYEVRRTFNNDMLLAFISSCIAALVYILTSCVFLSFFGI-----ASIGLSCLVALF 199

677 LGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGNADNIFIVLEQRLPR-RPGEPREVH 735
200 LYHVVFGEI---QYLGILG-----VAAFVIVIGVDDVFFVINTYRQATHLEDQRLMIH 251

736 IGRALGRVAPSMLLCSLSEAICFFLGTALTPMPAVRTFALTSLGLAVI----LDLLOMSAFV 792
252 TVQTAGK---ATFTSLTTAAAYAAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308

793 ALLSLDSKQEQASRLDVC--CCVKQELPPPG 822
309 GLWSLYLAPLESSCQTSCHQNCRSKRTSLHFP 340

Query Match 2.0%; Score 141.5; DB 6; Length 1048;
Best Local Similarity 18.1%; Pred. No. 0.011;
Matches 145; Conservative 110; Mismatches 243; Indels 303; Gaps 32;

624 EINRTTADLPFATSYIVIFYISLALGSYSSWRVMDSKATL--GLGGVAVVLG--A 679
339 EVVKTLE-----AIIIVFLVMYFLQNF-----RATLIPFIAPVVLGTPA 381

680 VMAAGFFSYLGIRSSVLQVVPFLVLSVG--ADNIFIVLEQRLPRRPGEPREVHIG 737
382 VLAAGFP-----SINTLTMFG-WVLAIGLVDDAIVVENVERVMTTEGLPPKEATR 432

738 RALGR-----VAPSMLLCSLSEAICFFLGTALTPMPAVRTFALTSLGLAVILDLLOMSAFV 792
433 KSMGQIQGALVGIAMVLSAVFIPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALITP 490

793 ALLSLDSKQEQASRLDVCCKVKQELPPPGQEGLLGFPQKAYAPFLLHWI----- 844
491 ALCA-----TMLAPVAKGDHGEKGFHFNRLFDKSTHYTDSVGNILR 536

845 TRGVVLLFLALF-----GVLSYSMCHISVGLDQELALPKDSYLL 884
537 STGRYLLLYLIIVVGWYLVFLPSSFLPDEQGVFL-TMVQLPAGATQERTOKVLDEVT 595

885 DYFL-----FLNR----- 892
596 DYVLNKEKANVSFVAVNGFAGRGQNTGIAFVSLKDNADRGKKNKEVATQRTATAAF 655

893 ----YFEVGAPVYFTTIGVNFSSBAGNNAICSSAGCNFSFTQ----- 932
656 SQIKDAMVFAFNLPAIVLGTATGDFE-----LIDQAGLGHKLTOARNQLFGVAK 708

933 ----KIYATEPEQSYLAIPAS-----SWDDFDIWLTP 963
709 YPDLIVGVPRNGLEDTPQFKIDIDQEAQALGVSISDINTTLGAAMGGSYVNDIFD---R 765

964 SSCCRLYISGPNKDKFCPTVNSINCLKNCSITMGSVRPSVQFHKYLPWELNDRPNIK 1023
766 GRVKVYVWSEAKYRMLPDDIND-----WYVSGSGQM 798

1024 CPKGLAAVST-----SVNLTSDGOVLSRGMAYHKPKLNQSDQYTEALRAARELANIT 1077
799 VP----FSAFSSRWYSGSPRLRYNGLPSEILGOAPGKST---GEAMAMEELASKLP 852

US-11-242-459-34
; Sequence 34, Application US/11242459
; Publication No. US20060035835A1
; GENERAL INFORMATION:
; APPLICANT: KITA, Shunbo
; APPLICANT: TANIMAWA, Yoshio
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703US0P
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 34
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-34

Query Match 2.1%; Score 143; DB 7; Length 456;
Best Local Similarity 25.4%; Pred. No. 0.0025;
Matches 69; Conservative 41; Mismatches 106; Indels 56; Gaps 12;

576 LNNYPAGDPRLAQAKLWEEAFLEEMRAFORRMAGMFQVFTFAERSLED----- 623
100 LPNYSVDDR-----WEE-----QRAKFSFVTVVAMLAQKSTSKVQVLYGGTD 144

624 ----EINRTTAD--LPFATSYIVIFYISLALGSYSSWRVMDSKATLGLGG-VAVV 676
145 LPDYEVRRTFNNDMLLAFISSCIAALVYILTSCVFLSFFGI-----ASIGLSCLVALF 199

677 LGAVMAAGFFSYLGIRSSVLQVVPFLVLSVGNADNIFIVLEQRLPR-RPGEPREVH 735
200 LYHVVFGEI---QYLGILG-----VAAFVIVIGVDDVFFVINTYRQATHLEDQRLMIH 251

736 IGRALGRVAPSMLLCSLSEAICFFLGTALTPMPAVRTFALTSLGLAVI----LDLLOMSAFV 792
252 TVQTAGK---ATFTSLTTAAAYAAANVFSQIPAVHDFGLFMSLIVSCCWLAVLVTMPAAL 308

QY 1078 ADLRKVPCTDPAPEVFPYTTNNFYEQYLT-----ILPGLFMLSICL----- 1120
Db 853 SGI-----GYDWTGMSYQBELSGNQAPALYAISLIVVFLCLAALYESWSIPF 899
QY 1121 -----VPTFAVSCILLGLDLRSGLNLLSVMLIVDTVGFMAWDISYNAVSLIN----- 1170
Db 900 SVMVLVPLGVIGALLAA--TFRGLTNDVYFQVGLTTIGLSA-----KNAILIVEFAKOL 952
QY 1171 -----LVSAGVMSVEFVSHITRSFAISTKPTWLERAKEATISMSGSAVPAGVAMT 1219
Db 953 MDKEGKGLVEAMLEAVMRRLRPILMTSLAFMLGVMP-----LVSSG-----AGSGAQ 1000
QY 1220 NLPGILVL-GLAKAQLIQIF 1239
Db 1001 NAVGTGLGVMTATVLAIF 1021

RESULT 11
US-10-392-234A-16
; Sequence 16, Application US/10392234A
; Publication No. US20050255538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; PRIOR FILING DATE: 2003-03-17
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 1046
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-392-234A-16

Query Match 2.0%; Score 140.5; DB 6; Length 1046;
Best Local Similarity 20.5%; Pred. No. 0.013;
Matches 187; Conservative 126; Mismatches 305; Indels 293; Gaps 49;

QY 506 NOTLMGOT--SQVDWKDHFYLCANAPLTFKDGFTALALSCWADYGAPVFPFLAIGYKGD 563
Db 231 NATIIGKTRLOTAEQENILLKVN-----PDGSOVRLKVDVVG-----LG----GQD 274
QY 564 YSAREALIMTFSLNYP-----GDPRLAQAKLWEE--APLEEMRAFORMAGMFQ 612
Db 275 YS-----INAPNGSPASGAIKATGANALDTAKAIRQTIANLEPFMPQGMKVVPYD 328
QY 613 VFTAERSLEDEINRTAEDLPATSYIVIFLYISLALGSSYSWSRVWDSKATL--GL 670
Db 329 TTPVWSASIH-EVVKTLGE-----AIIIVFLVWYFLQNF-----RATLIPTI 370
QY 671 GGVAVVLG--AVMAAGFFSYLGIRSSVLTLQVVPFLVLSVG--ADNIFVLEVORLPR 726
Db 371 AVPVVLGTFGVLAAGF-----SINTLTFG-MVLAIGLLVDDAIVVENVERVMA 421
QY 727 RFG-EPREV---HIGRALGR-VAPSMLLCSLSAICFFLGCALTPMPAVRTFALTSLGLAVI 781
Db 422 EEGLSPREAARKSMGQIQALVGIANVLSAVFLPMAFFGSGTGV--YQFSITIVSAMA 479
QY 782 LDFLLQMSAFVALLSDSKRQESRLDVCCC---VKPQLPPPGQEGEGLLLGFFOKAYAP 838
Db 480 L-----SVIVALI-----LTPALCATMLKPIEKGDGHEKGFQGFENRMLFS 522
QY 839 FLLHWITRGVV-----LILLFLALFGVLSYMSCHISVGL-----DQ-----ELALPKD 880
Db 523 -TTHGYERGVAStLKHRAPYLLIYVIVAGMIWMFTRIPTAFLPDEDOGVFAQVOTPPG 581

QY 881 S-----YLLDYFLFLNRYFEVGAAPVYFVTTILGYNFSSE----- 913
Db 582 SSAERTQVVVDSSMREYLLK-----ESSSSSVFTVTGFNFAGRGQSSGMAFIMLKP 633
QY 914 -----AGMNAI-----CSSAGCWNFSFTQIOYATEFFEQSYLAIPASSWVDDFD----- 959
Db 634 WEBRPGGNSVFLAKRAQMHFFSFKDAMVFA--FAPPSVLELGNATGFDLFLQDQAGVG 691
QY 960 -----WLTPSCCRLYISGPNKDKPCFSTVN-----SLNCLKNQMSI 996
Db 692 HEVLLQARNKFLMLAQNAPALQVRNMGMSDEPOYKLEIDDEKASALGVSLADINSTVSI 751
QY 997 TMGSV-----RPSVEQFHKYLPWFINDRPN-----IKCPKGLA---AYST 1034
Db 752 AMGSSVYVNDPIDRGRVKRV-----YLOGRPDARMNPDDLKSWYVRNDKGEWVPFNAPAT 805
QY 1035 -----SVNLTSDGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDP 1088
Db 806 GKWEYSGPKLERYNGVPAMEILGEPAPGLSSGD---AMAAVEE---IVKQLPKGVG--- 855
QY 1089 APEVFPYTTITNVFYEQYLT-----ILPEGLFMLSIC-----LVPTFAVS 1127
Db 856 -----YSWTGLSYERLSGSQAPALYALSLVVFLCLAALYESWSIPFSVMLVPLGVI 909
QY 1128 CILLGLDLRSLNLLSIVMLIVDTVGFMAWDISYNAVSLINL-----V 1172
Db 910 GALLATSMR-GLSNDVFFQVGLTTIGLSA-----KNAILIVEFAKELHEQKGIWEAAI 963
QY 1173 SAVGMSVEFVSHITRSFAISTKPTWLERAKEATISMSGSAVPAGVAMTNLPGILVL-GLAK 1231
Db 964 EACRMELRPVMTSLAFILGVVP-----LAISTG-----AGSGSQAHTGTGVIGMVT 1011
QY 1232 AOLIQIIFPFR 1242
Db 1012 ATVALIFWVPL 1022

RESULT 12
US-10-467-657-2648
; Sequence 2648, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 2648
; LENGTH: 1067
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-2648

Query Match 1.9%; Score 134; DB 6; Length 1067;
Best Local Similarity 19.4%; Pred. No. 0.043;
Matches 185; Conservative 139; Mismatches 354; Indels 274; Gaps 44;

QY 452 RHLQWV-SPEAQR--NISLQIDYCAFLPNPDNTSLYDCCINSLLQVFPNNRTLLLTANOT 508
Db 180 RAMRIVDPKLLQNYNLSPADVGA-LSAQNIQISAGSIGSLPAV--RQGT---VTATVT 233
QY 509 LMGTQSDVDKDHFLYLCANAPLTFKDGFTALALSCWADYGAPVFPFLAIGYKGYSEAE 568
Db 234 AOGQLGTAEFGNVLIRANT-----DGSNIYLDKVAKVGLGM-----EDYSSST 277

GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ. ID NOS: 59
; SEQ ID NO 16
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-16

Query Match 1.8%; Score 126.5; DB 7; Length 200;
Best Local Similarity 25.7%; Pred. No. 0.015;
Matches 53; Conservative 38; Mismatches 92; Indels 23; Gaps 8;

QY 612 QVTFTAEKSLDRIINTTAD--LPIPATSYIVIFLYISLALGSYSGRWVVDKATLG 669
DB 3 QVLYGGTDLFDEYEVRETFNDMLLAFISSCIAALVYILTSCSVFLSPFGI-----ASIG 57
QY 670 LGG-VAVVLGAVMAAGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLEYQRLPR-R 727
DB 58 LSLCLVALFLYHVFGI---QYLGLNG-----VAAPVIVGIGVDDVVFVINTYRQATHLE 109
QY 728 PGEPREVHIGRALGRVAPSMLLCSLSEALCFILGALTMPMPAVRTALTSLGLAVI---LDF 784
DB 110 DPQLRMHTVQTAGK---ATFFTSLTAAAYANVFSQIPAVHDFGLFMSLVSCCWLAV 166
QY 785 LIOMSAFVALLSLDSKQASRLDVC 810
DB 167 LVTMPAALGLWSLYLAPLESSCQTC 192

RESULT 15
US-10-793-626-2020
; Sequence 2020, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2020
; LENGTH: 801
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: amino acid sequence
US-10-793-626-2020

Query Match 1.8%; Score 126; DB 6; Length 801;
Best Local Similarity 19.9%; Pred. No. 0.12;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;
QY 841 LHWTR-----GVVLLFLALFGVY-----LYSM-----CHISVGLDQ 873
DB 62 LNMHPHGMFNLYVDGLLFLSLTIGSLVLSYGLSKSEQLGNFYCYLLLFMGA 121

QY 874 ELALP-KDSYLLDY-----FLPLNRYFEGAPVY-----FVTTLTGYNPSSBAGM 916
DB 122 MLGVLSDFNFIILYFELWELTSFSSFLILISFWREKASIIYGAQKSLIITVLG-GLSMLGGI 180
QY 917 NAICSSAGCNPSTQKIQVATEPPEQSYL-----AIPASSWDDFDIDW 960
DB 181 --LLSLATDTFSIQAMISKASDIQNSPPFIFLWILFMIGAFTKSAQVPFYIWLPDAMEA 238
QY 961 LTPSSC-----CRLYISGNKDKFCPS-----TVNSLNCILKNCMSITMGSVRPSVE 1006
DB 239 PTFVSAYLHSATVMKAGLYLIARITIPALISEGWVITIT-----LVGLITILFWASLNATKQ 294
QY 1007 QFHKYLPWFNLDRPNIKCPKGGAAYSTVNLFSDGOVLASRFMAYHKPLKNSQDYTEAL 1066
DB 295 HDLK-----GILAFSTVSQLGIMSMGLGIGAVSYHYQGANSQLYVAGF 337
QY 1067 RAARELAANI-----TADLRKVPGETDPAFEVPPYITNTVFEYBYLT 1107
DB 338 VAAIFHLINHATPKGALFMITGGIDHSTGTRDKLGG-----LLTIMPISFT-----LT 387
QY 1108 ILPEGLFMLSCLIVPTFAVSCLLILGLDLRSGLINLLSIVMILVDTVGFMALWDISYNAVS 1167
DB 388 VIT-----TISMACVPPF--NGFLSKKEKFLSMINVTILNLSLNTLG-----LLPIIA 435
QY 1168 LINLVSAGMSVEFVSHITRSFAISTKPTWL-ERAKEATISMGSAVPAGVAMTNLPGILV 1226
DB 436 IIGSIPTFVYSIKFILHI---PFGSYKPEALPKQAHESILM---LISPIILTSL--VIV 487
QY 1227 LGAKAQLIQIFPRLNLLITLLG-----LHGLVFLPVILSYVG 1266
DB 488 FGLFPSILTQSIIBPASVAVSOTSNTABPHLPHGKIT--PAFLSTIG 532

RESULT 16
US-10-392-234A-34
; Sequence 34, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acrab Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 34
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-392-234A-34

Query Match 1.8%; Score 121.5; DB 6; Length 1043;
Best Local Similarity 17.7%; Pred. No. 0.4;
Matches 236; Conservative 166; Mismatches 398; Indels 531; Gaps 58;

QY 134 SLFINVTR---VAQLGAGQLPAV---VAYEAFYQHSFAEQSYDSRVRVPAAATLAVG 186
DB 16 ALPISLAGLLVISKLPAQYPNVAPPOITITATYP-----GASAKVLD 59
QY 187 TWCYVYGSALCNAQRLNFGQ-DTNGGLAPLDITFHLLEPG----- 226
DB 60 SVTSVLRESLNGAKGLLYPESTNNGTAGIIVTP---EPGTDPLDAQVDVQNLKKAEA 116
QY 227 ---QAV-MSGIQ-----PLNEGVARCNESQDDVATCSQDCAACPAIAR 268
DB 117 RMPQAVLTQGLVEQTSAGFLIYALSYPEGAQRSDTTALGDYAARNINNELRLRFGVGK 176

; TITLE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2923
; LENGTH: 914
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-072-512-2923

Query Match 1.7%; Score 115.5; DB 7; Length 914;
Best Local Similarity 18.8%; Pred. No. 1;
Matches 208; Conservative 140; Mismatches 381; Indels 375; Gaps 55;

QY 55 VSCLSNTPARKITG-----DHLIL-----LQICPRLYTPG-NTQACCSAKQLVS 98
DB 19 VICLPRPP--KVLGYRLEPPHLLTACTLEGMYNIEELERDINGTLLSQATCELCD--G 74
QY 99 LEASLSITKALLTRCPACSDNFVNLHCHNTCS-PN---QSLFINVT-----RVAQLGAG 148
DB 75 NENFMVNVNAGDCRCVCEPTFVNTSRSCACSEPNILTGGLCFSGTGNFRLRISAARYG 134
QY 149 QLPVAVVEAFYQHSFABQSYDSCSRVRVPAATLAVGTMG-----GVYGSALCNAQRWLN 204
DB 135 EVGNSLTSEWFAKYL---QSSAAACWVYANLTSCQALGSMCMNMNSYDFATFDACGLFQ 191
QY 205 FQDGTGNGIAPLDTIFHLE-----PGQAVGSGIOPLNEGVARCNEQSGDDVATCSCQD 258
DB 192 FIFENTAGLS-----FVHSISFWRQMLPWFYGDQL-----GLA----- 225
QY 259 CAASCAPAIARPAQDSTFYLGQMPGSLVLIILCSVFVAVVTILLVGFVRVAPARDKSKVD 318
DB 226 -----FQVLSST-----SLPTNFSF----- 240
QY 319 PKGTSLSKLSF--STHTLQPFQCGWGTWVASHPLTILVLSVIPVVALAAGLVFTLT 376
DB 241 --KGEDQNTKLFVAASYDIRGNFLK-WQT-----LEGV----- 272
QY 377 TDPVELSNAPNSQASEKAFHQHFGPPPTN-----QVILTAPNRSSRYRDSLLGPK 430
DB 273 -----LQCPDTEITLNNAY---SGTTYQNCPEIPISKILIDFPPIFY----- 314
QY 431 NFGSILDLDLLELLELOERLRHLQWSPQARNISLQDIYAPLNPDPNTSLYDCCINSL 490
DB 315 -----DVLEYTDENHQVILAV-----PVLNMLNH-----NKI 344
QY 491 LQYFQNNRLLLTANQTLMGQTSQYDWDKHELYCANAPLTKDGTALASCSMA---DYG 547
DB 345 FVNQDSNGKWLITRIFLV---DAVSGRENDL---GTQPRVIRVATQISLVSVHLVPNTIN 399
QY 548 APVFPFLAIGGKGYKDYSEAE-----ALIMTSLNNYPAG--DPLRAQAKLWEAFLEEM 600
DB 400 GNTYPELLITITAYSDIDIKNANSQSVSVSVTYEMDGEAHVQTDIALGVGLAVLASL 459
QY 601 --RAFORMAGMFQVTTAERSLEDRINRTAEDLPFIPATSYIVIFLYISLGSYSWS 658
DB 460 LKTAGWKRGISPMIDL-----QTVWKFVYVYAGDLANVFPIITVGTGLY--W- 505
QY 659 RMVVDKATLGLGGVAVVLGAVMAAGPFYSYLGRSLVILQVVPFLVLSVGADNIFIV 718
DB 506 --LIFFAQ---KSVSVLLPMPQOERFVTVGCAFKALQFLHKLISQITID---VFF 557
QY 719 LEYQRLPRRPGEPREHVHIGRALGRVAPSMLLCSISEAICFFLGTALTPMAVRTF----- 772
DB 558 IDWER-----PK-----GKVLKAV-----EGEGGVSRSATVPVSIWRTYFVANEW 596
QY 773 -----ALTSGLAVILDLLQMSAFVALLSLS-----KRQEASRLDVCCKVQPEL 818

QY 738 RALGR-----VAPSMLLCSLSEACIFFLGAITPMPAVRTFALTSGLAVIDFLQMSAFV 792
DB 433 KSMQIQOALGVIANVLSAVFVPAFPGSGTGAII--YRQFSITIVSAMALSVLVALITP 490
QY 793 ALLSLSKROBASRLDVCCCVKPOELPPPGGEGILLGFPQKAYAPFLLHWI----- 844
DB 491 ALCA-----TMLKPIAKGDHGGKKGFGFWFNRMFEKSTHHTYDVSVGILR 536
QY 845 TRGVVLLFLALFGVSLY-----SMCHISVGLDQE----- 874
DB 537 STGRYLYLYLIIVGMAYLFLVRLSSFLPDDQGVFMVWQLPAGATQERTQKVLNEVTH 596
QY 875 --LALPKDS-----YLLDYFLFLNR----- 892
DB 597 YLITKEKNVSVFPAVNGFGAGRGQNTGIAPVSLKDWADRPGEENKVEAITMRATRAFS 656
QY 893 -----YFEVGAPVYFVTTLGYNFS--SEAGM-----NAICSSAGCANNPSFT-- 931
DB 657 QIKDAMVAFNLPALVELGTATGDFDELIDQAGLGEKLTQARNQLLAENAKHPDMLTSV 716
QY 932 -----QKIQYATEPEQSYLAIPAS-----SWDDDFIDWLTSPSSCCRLYI 971
DB 717 RPNGLDTPQFKIDIDQEKAAQALGVSIINDINTLGAANGSGSYNDVID---RGRVKKYV 773
QY 972 SGPNDKFCPSTVNSLNCNLKNCMSITMGSVRPSVEQFHKLFPWFLNDRPNIKCPKGLAA 1031
DB 774 MSEAKTRMLPDDIGD-----WYVRAADGQWVP--FSA 803
QY 1032 YST-----SVNLTSDQVLSRFMAHYHKLKNSQDYETALRAARELANITADLRKVP 1085
DB 804 FSSSRWEYSGPRLERYNGLPSMEILLQQAAPGKST---GEAMELMEQLASKLPTGV----- 855
QY 1086 TDPAEVFPYTTITNVEYQYL-----TILPGLFMLSCL-----VPT 1123
DB 856 -----GYDTMGMSQBERLSGNQAPSLYATSLIVVFLCLALYESWSIPFSVMLVPL 907
QY 1124 FAVSCLLGLDLSRGLNLLSIVMILVDTVGFMAIWDISYNAVSLNL----- 1171
DB 908 GVIGALLAA--TPRGLTNDVYFQVGLTITIGLSA-----KNAILIVEFAKDLMDKEGKL 960
QY 1172 ----VSAGVMSVEFVSHITRSPAISTKPTWLERAKEATISMSGAVPAGVAMTNLPGLVL 1227
DB 961 IEATLDVAVRMLRPILMTSLAFTILGVMP-----LVISTG---AGSGAQNAGVTGM 1008
QY 1228 -GLAKAQLQIIF 1239
DB 1009 GGMVTATVLAIFF 1021

RESULT 20
US-11-072-512-2923
; Sequence 2923, Application US/11072512
; Publication No. US20060029945A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAOHICO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOTUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO

```
Db 597 NEIQVRKINSFLQVLTVL--PFEVVGFKNLALMDSSSLRNPFSYIAPYSCI----- 649
Qy 819 PPRQCEGLLLGFQKAYAPFLHWTGRVVLVLLFLALF-----GVLSYMSCHI 867
Db 650 -----IRYAVSAALMLAIGIIQVFFAVFERIEDKIQFVDCLSMSNI 694
Qy 868 SVGLDQELALPKDSYLLDYFLPNRYFEVGAPYV-----FVVTLYGNFSSSEAGMNAICSS 922
Db 695 SV-----FLLSHKCF--GYVIHGRSVHGHADTNMEEMNNMLKREA--ENLCSQ 738
Qy 923 AGC--NNFSFTQKIQVATEPEOSYLAIPASSWVDFFIDWLTSSCCRLYISGNKDKFC 980
Db 739 RGLVPNTDQGTFAIASNQM--RQHYDRIHET-----LIRKNGPARLLSSS----- 782
Qy 981 PSTVNSLNCNKMSITMSVRSVEQFKYLPWFNLNDRNPKCPKGLAAAYSTSVNLT 1040
Db 783 ASTFEOSIKAYHNMFLGSP----IDHVHKEMDYFIKDLLE----- 822
Qy 1041 DGOVLASRFMAYHKPLKNSQDYTE 1064
Db 823 --RILGWEMF--EPWEKSIFYND 841

RESULT 21
US-10-392-234A-30
; Sequence 30, Application US/10392234A
; Publication No. US2005025538A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazhi Li
; TITLE OF INVENTION: Method for Screening for acRAB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 30
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-392-234A-30

Query Match 1.6%; Score 113; DB 6; Length 1034;
Best Local Similarity 18.8%; Pred. No. 1.9;
Matches 161; Conservative 115; Mismatches 270; Indels 312; Gaps 38;

Qy 576 LNNYPA-----GDPRLAQNKLWEAFLEEMRAFORMAGMFOVTTAERSLE-DEI 625
Db 281 INKPAAGIGIKATGANALDTAKAKLAELQPFPPQMKVLYPYDTTPFVQLSIHEV 340
Qy 626 NRTTARDLPFATSYIVIFLYLSLALGYSYSSRVNVDGKATI--GLGGVAVVLG--AVM 681
Db 341 VKTLFPAIMLV---FLVNYLFQ-----NRRATLIPTIAPVVLGTFALL 383
Qy 682 AAMGFFSYLGRSSLVILQVPELVLSVG--ADNIFVLEYOR-----LPRRGPPEPR 733
Db 384 AAFGY-----SINTLWFG--MVLAIGLLVDDAIVVENVERVMEDKLP--PKEATE 432
Qy 734 VHGIRALGR--VAPSMCLCSLEAICFELGALTMPAPVTFALTSLGLAVILDFLLQMSAFV 792
Db 433 KMSQIQGALVGAMVLSAVFIPMAFFGGSTGAI--YRQFSITIVSAMALSVLVALILTP 490
Qy 793 ALLSLDSKQEARSLDCCVCKPQELPPCQGGGLLGFQKAYAPFLHWM----- 843
Db 491 ALCATLLKPVSAEH-----HENKGGFGFWNTTFDHSVNHYNSTNSVKILG 535
Qy 844 -----ITRGVVLVLLFLAL-----FGVLSYMSCHISVGLDQELALPKDSYL 883
```

```
Db 536 STGRYLLIYALIVAGNV--VLFLRLPSSFLPEEDQGVFL-TMIQLPAGATQERTQKVLQDV 593
Qy 884 LDYFLFLNRYFEVGAPYVVTTL--GYNFSSEAGMNA----- 918
Db 594 TDYLLKNEK-----ANVESVFTVNGFSPGQA--QNAQMAFVSLKPMWEERNGDENSABAVI 647
Qy 919 -----ICSSAGCANNFSFTQ----- 932
Db 648 HRAKMGELKIRDOGFVTPFNMPAIVELGTATGDFDELIDQAGLGHDLALTOARNQLLGWAAQ 707
Qy 933 -----KIQVATEPEOSYLAIPAS-----SWDDFDIDLMTLP 963
Db 708 HPASLVSVPNGLEDTAQFKLEVDEQKAQALGVSLSDINOTISTALGGTYVNDFID--R 764
Qy 964 SSCCRILYISGNPKDKFCPSTVNSLNCNKMSITMSVRSVEQFKYLPWFNLNDRPNIK 1023
Db 765 GRVKUYQADAKFRMLPEDV-----DKUYVR 791
Qy 1024 CPRGGLAAYSTSVNLTSDQVTLASRFMAYHKPLKNSQDYTEALRA-----ARELAANITA 1078
Db 792 SANGENVPFSA---FTTSHWYVGSPLERYNGLPMSMEIQGEAAAPGTSSGDAMALMENLAS 848
Qy 1079 DLKRVGTPDAFEVFPYTTITNVFYEQYLT-----ILPEGLFMLSCLVPTFF-----AVS 1127
Db 849 KLPAGIG-----YDWTGMSYQERLSCNQAPALVAISFVVVFLCLAALYESWSIPVS 899
Qy 1128 CLL-----LGLDLRSGLNLLSIYMLV---DTUGFMAWDISYNAVSLINL----- 1171
Db 900 VMLVPLGIVGVLAAATLNFQKNDVYFMVGLLTTIGLSA-----KNAILIVEFAKOLMEK 954
Qy 1172 -----VSAGMSVSEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLP 1222
Db 955 EKGGVVEATLMAVRWELRILMTSLAIFLGVL-----LAISNG-----AGSGAQNVA 1002
Qy 1223 GILVL-GLAKAQLIQIFF 1239
Db 1003 GIGWGMVSAITLLAIFP 1020

RESULT 22
US-11-082-389-362
; Sequence 362, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
```



```
Qy 1067 RAARELANITADLRKVPDPAPEVPPYTTTNYFEOYLTILPEGLFMLSCLVPTFAV 1126
Db 476 RSTTHIAN-----AYGIAEV-----GYMMVSTVLV----- 500
Qy 1127 SCLLLGLDLRSLNLLSIVMLV-DTVGFMAL-WDISYNAVSLINLSVAGMGVVE--FV 1182
Db 501 -----TLVMLLWQTNIFALCFPLIFGSEVETIYLLAVTKLLEGGWV 543
Qy 1183 SHITRSFAISTKPTW-----LERAKEA--TISM-----GSAVFAGVAMTNLPGLVLGL 1239
Db 544 PLVATPFLTWYIWNYSVLKYQSEVRERISMDFMRELST-----LGTIRIPGI---GL 596
Qy 1230 AKAOILQ-----IPFRLLNLTLLGLLGLVFLPVLVSYGPDVNP 1271
Db 597 LYNELVQGISIFG---QFLTLTPAIHSTIIF--VCIKYVPVPVVP 637

RESULT 26
US-11-098-686-11239
; Sequence 11239, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11239
; LENGTH: 1049
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-11239

Query Match 1.6%; Score 109.5; DB 7; Length 1049;
Best Local Similarity 21.1%; Pred. No. 3.6; Indels 126; Gaps 19;
Matches 72; Conservative 68; Mismatches 126; Indels 75; Gaps 19;

Qy 960 WLTPSSCCRLYISGPNKDFCPSTVNSLNCNCSITMGSVRPSVEQFHKL--PWFLN 1017
Db 198 WLKPRMEQLKIT-PN-----DVINAI---NDQNIQIGAGRIGDEPVSSSIGVTWQII 246
Qy 1018 DRPNIKCPKGLAAYSTSVNLTG-DQOVLASRPM----- 1051
Db 247 TKGRLETP-----QFENIILRTPDGSILRLKDIARIELGAQVYNFTGRQDGLAAIPIGI 302
Qy 1052 YHKPLKNSQDYTEALRA-ARELAANITADLRKVPDPAPEVPPY---TITNVFYOYLT 1107
Db 303 FLAPGANALTAEAVRVKMKELSTQFPVG-----AYDI--PYDTTFVNIKEVK 353
Qy 1108 ILPEGLFMLSCLVPTFAVSCLLIGDLRSLNLLSIVMLVDTYVGFMAWDISYNAV 1167
Db 354 TLCEAMFLV-----FLVYVFL-QNWRATLPLCLAVPVSIYGTAGMHLGFSINTLT 405
Qy 1168 LINLSVAGMSVE-----FVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAM---TN 1220
Db 406 MFLVLVLAIGIVDDAIVULENVERHIS-EGQPVHIATAK-AMSEVTSATIIAIVLVCVAF 463
Qy 1221 LPGILVGLAKAQILQIIFFRLLNLTLLGLLGLVFLPVI 1261
Db 464 IPVAFIGGLA-GRMYQOQFAITIAVSVISGVV-ALTFTPAL 502

RESULT 27
US-10-392-234A-67
; Sequence 67, Application US/10392234A
; Publication No. US2005025538A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Pharmacia and Upjohn Corporation
; APPLICANT: Buxser, Steven
; APPLICANT: Poole, Keith
; APPLICANT: Decker, Douglas
; APPLICANT: Xiazanhi Li
; TITLE OF INVENTION: Method for Screening for acraB Transporter Family Inhibitors
; FILE REFERENCE: 6206
; CURRENT APPLICATION NUMBER: US/10/392,234A
; CURRENT FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: US 60/364,935
; PRIOR FILING DATE: 2002-03-15
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 1032
; TYPE: PRT
; ORGANISM: Haemophilus influenzae
US-10-392-234A-67

Query Match 1.6%; Score 109; DB 6; Length 1032;
Best Local Similarity 18.5%; Pred. No. 3.9; Indels 232; Gaps 38;
Matches 147; Conservative 129; Mismatches 285; Indels 232; Gaps 38;

Qy 625 INRTAEDLPFATSYIVIFLYISLALGSSYSGRVMVDSKA-TLGLGGVAVVLG---- 678
Db 331 INSSHEVIKTIGEATLVLVILMFIGSFA---ILIPILAIPISLIGVLMLLQSFNFS 387
Qy 679 -----AVMAAMGFFSYLGRSSRSLVILQV-----PFLVLSVAGADNIFIVLEVQ 722
Db 388 INLMTLLALILAIG---LVVDDAIVLENDIRHIKAGETFFRAAIIGTREIAVFPV---- 439
Qy 723 RLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFFGLALTPMPAVTFAITGLTGLAVIL 782
Db 440 -----ISMTIALIAYVSPNAL-----MGITG-TLPKEFALTLAGAVFI 477
Qy 783 DFLQLMSAFVALL---SLDSKRQEAERLDVCCVXPQLPPPGOGEGLLGFFQKAYAPF 839
Db 478 -----SGVVALTLSPMSSKLLKSN-----AKPTWMEERVEH---TLGKVRVVEYM 521
Qy 840 L-LHWITRGVLLFLALFGVSLYSVMCHISVGLDQELALPKD-----SYLLDY 886
Db 522 LDVLMLNRKSLMAFAVWIFS---TLPLFNLSELSTPNEDKGAFIAGNAPSSVNVVDY 577
Qy 887 FL-----FLNRVFE-----VGAPVYFVTTLGYNPSSEAGMNAIC-----S 921
Db 578 IQNAMPQYMKVNYMETPEVSFGMSIAGAP-----TSNSSLNIIITLKWKERSRQ 626
Qy 922 SAGCNFSTQKIQVATEPPEQSILAI-----PASSWVDDFDIMLTPSSCCRL 969
Db 627 SAIMNEIN-----EKAKSIPEVSVAFNIPEDTGEQGPVSVILVKTADYKSLANTABK 681
Qy 970 YISGPN-KDKFCPESTVNSLNCNCSITM-----GSVRPSVEQFHKLPHFLNDRPN 1021
Db 682 FLSAMKASGKF---IYNLDLTYDTAQMTISVDKEKAGTYGITMQQISNTLGSFLSGATV 738
Qy 1022 IKCPKGLAAYSTSVNLTGQOVLASRFMAVYHKPLKNSQDYTEALRAARELAANITADLR 1081
Db 739 TRVDVVG-RAYKVISQVRKDRDLSPESFONYIYLTASNGQSV-----PLSSVISMKLE 789
Qy 1082 KVPGTDPAF-----RVFFPYTITNV-----FYEQYLTILPEG-----LFML 1116
Db 790 TQTSIPRPSQLNSAISAVPMPISSGDAIWLQQQATDNLPGQYTFDPKSEARQLVQE 849
Qy 1117 SLCLVPTFAVSCLLLGL-----DLRSLGNLLSI-----VMILVDTYVGFMAWDISY 1163
Db 850 GNALAVTFALAVIIIFLVAIQPESIRDPVMIVISVPLAVSGALVSLNLSFPIAGTTL 909
Qy 1164 NAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGV---AMTN 1220
Db 910 NIYSQVGLITVLGL-----ITKHGIL-----MCEVAKEEQLNHGKTRTEATHAAKVR 957
Qy 1221 LPGILVGLAK-AQLIQIIF-----FRLLNLTLLGLLGLVLF-----LPVILSVG 1266
```

Db 958 LRPIIMTTAAWAGLIPLLYATGAGVRSFISGIVI--VAGLSIGTITFTLFLVLPVVSIVA 1016
QY 1267 PDVNPALALEQXR 1279
Db 1017 TEHKPLPVFDENK 1029

RESULT 28
US-11-087-099-4135
; Sequence 4135, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4135
; LENGTH: 648
; TYPE: PRT
; ORGANISM: Sulfolobus tokodaii
US-11-087-099-4135

Query Match 1.6%; Score 108.5; DB 7; Length 648;
Best Local Similarity 19.5%; Pred. No. 2.2;
Matches 136; Conservative 99; Mismatches 249; Indels 213; Gaps 37;

QY 604 QRMAGMFOVTFATERSLEDEINRTTAEDLPATSYIVIFLYISLALGSYSWSRWVD 663
Db 3 ERETKARDIGISDRQLRRSLGKFE-----LLYLSLGGIIGSGW---LFA 45
QY 664 SKATLGLGGVAVVLGAVMAAMGFSSYLGRSSVLQVVPFLVLSVGCADNIPFVLEYQR 723
Db 46 SLATAYAGASAILSWIAG-----ILVNFICLVAEIGA-----A 81
QY 724 LPRRPGEPREVHI--GRALGRVAPSMLLCSLSAICFFLGAITPMPAVRTFALTSLGLAVI 781
Db 82 IPKSGGIGTRYPHYTHGGLVGYLI-----TWAYFLSAAS-VPAIRAAAIEIGSY 130
QY 782 LDFELLQMSAFVALLSDSKRQASRLDVCCVQBELPPQEGE---LLGFQKAYAP 838
Db 131 YPOLITGTGF-----DGTVTI-----LPLGIGLAGLLILFPFLNFGV 171
QY 839 FLIHMITRIG-----VLLLFLLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFN 891
Db 172 NILGKVTHGAGWKLLIPTITFLLALLDLHS--ANFSLGG---FLPSAEYVKG---GSS 224
QY 892 RYFEGAPVYFVTLGYNFSSEAGMNAICSSACNNFSTQKIQYATEPPEQSYLAIPAS 951
Db 225 GIYGFSAVLVAIPSTGVIFS-----YLGFRQAVEYGBE-----257
QY 952 SWDDFIDMLTPSSCCRLYISGNKDKFSTVNSLNCMLKMSITMGVSRSVEQ--FH 1009
Db 258 -----GNPKKIDIPFAVIG-----SLLIAIVLYTLQVAFV 288
QY 1010 KYLPW---FLNDREN-1KCPKGLAAVSTSVNLTSDG-QVLASRFMAYHKPLKNSQDYTE 1064
Db 289 GGIDWNKVLNESGKLVPTVPGNWSALSTAV--TASGVPIAGFPVLVLSRLASVS-----341
QY 1065 ALRAA---RELANITADLRKVP-GTDPAPFVFPYTTINVFY----BOYLTILPEGLFMLS 1117
Db 342 GLAAAFPTALAVLLTIDAVVSPGTG---WIYGTSTRTLYAFASNGY---LPEIFLKIG 395
QY 1118 LCLVPTFA-VSCLLGLDLRSGLNLLSVMLVDVTFGMALWDISYNAVSLINLVSAGV 1176
Db 396 KTKIPTYSLIAALIIGF-----IFLLPPSWVALVGFISSTATVLTYYIMGGIG 442
QY 1177 MSV--EFVSHITRSPASTKPTWLERAKEATISMGSA-----VPAGVMTNLPFG 1223
Db 443 LTVLRKHAABLRNPFKLAAPAPI---TGAITATLAAGLIVYVSSFAVLFFYFTGVFL-GLFL 498

QY 1224 ILV-----LGLAKAQLIQIFFRFLNLLITL-LGLL 1252
Db 499 FFIFYADKMLGINK--VVSIIIVGVINLAINLGMGLL 532

RESULT 29
US-11-087-099-11457
; Sequence 11457, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11457
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293
US-11-087-099-11457

Query Match 1.5%; Score 106; DB 7; Length 443;
Best Local Similarity 24.3%; Pred. No. 2;
Matches 74; Conservative 43; Mismatches 102; Indels 86; Gaps 15;

QY 1076 ITADLRKVPQTPAPFVFPYTTINVFYEQVLTLPBGLFMLSCLVPTFAVSCLLGLDL 1135
Db 96 ITLDYLLIP-----VITLYLSASFQSLPAPVFWWILGIAI---FNTVINLIGIKV 145
QY 1136 RSG-----LLNLLISIVMI--LVDTVGFMAWDISYNAVSLIN-----LVSAGVMSVEF 1181
Db 146 VSKFSLWLLFGLQGVLLIAFIVGTIVANGTVSFNVSSFFYNASHFDLSGVVQATGIVI--203
QY 1182 VSHITRSP-AISTKPTWLERAKEATISMGSV-----FAGVAMTNLPG- 1223
Db 204 VSYL--GPDIAISTLS---BEANDPTKSGIKAVLSILIGSLFVIVTFFGFPVAPYSEL 258
QY 1224 -----ILVLG-LAKAQLIQIFFRFLNLLITLGLLHGLV-----FLPVILS 1263
Db 259 NPDTPALTITLGRVGSWLVHLAEFTLVLSFGLAGGEGQTAVSRILYSGMRDILPRLS 318
QY 1264 YVGPDPVNPALALEQKRAEBAVAAMVVASCPNHPRSVSTADNIYVNHSPSGSIKAGAIIN 1323
Db 319 YLHPRVYNTPW-----VAIILVGI-----VSVALSLSLSLTTSVNLISFGALFG 361
QY 1324 FLPN 1328
Db 362 FLSLN 366

RESULT 30
US-11-082-389-354
; Sequence 354, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompeius, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; FILE REFERENCE: BGI-131CPCN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09

```
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 354
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-354

Query Match          1.5%; Score 105.5; DB 7; Length 783;
Best Local Similarity 20.0%; Pred. No. 4.9;
Matches 145; Conservative 108; Mismatches 268; Indels 205; Gaps 34;

QY 628 TTAEDLPFATSVIVFLYISLALGSYSWSRVMVDSKATLGLGVAVVAVMAAMGPF 687
DB 194 TTSEIIGIGIAFTVLFTF-----GSLIAGLPLITAVIGVGIGALAILAT----- 240

QY 688 SYLGIRSSVILQVFPFLVLSVADNIFIFVL-----EYQRLPRRPEPREVHIGRALGR 742
DB 241 AFTDLNVPFLAVM--IGLAVGID-YALFILLSRYAEYKMER--ADAAGMAVGTAGSA 295

QY 743 V--APSMGLCSSEATCFPLGALTMPAVRTFALTSGLAVIDLDFLLQMAFAVALLSLDSK 800
DB 296 VFPAGATVIAIALVALIIDIGFLTAM-----GISAAFTVFVAVALIALTFFIALLGVFGG 349

QY 801 QEASRLDVCCCKPQELPPGOG-----EGLL-----LGFQKAYAPFLH 842
DB 350 HAFKGI-----PGIGNPPTKQWEOALNRRSKRSWKLQK--APGL-- 392

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVVF 902
DB 393 -----VAVVVLGLGALTIPAM-----NLQLSLPDS----- 419

QY 903 VTTLGNFSSEAGNNAICSSAGCNFSFTQKIQYATEFFEQSVLAIPASSWDDFDWLT 962
DB 420 TSNIDTTQROSADLMAEGFGAGYN-----APFLVIVDTHEVNADSTALQ 463

QY 963 PSSCCRLYISGPNKDKFCPTSVNSLCLKNCMSITMGSVRPSVQFHKYLPWFNDRPNI 1022
DB 464 P-----LIEAPEBEGEF-----DREQARFATYM--VTQTYNS 496

QY 1023 KCPKGGIAAYSTSVNLSDGQVLASRFMAYHKPL-KNSODYTBALRAARELAANITADLR 1081
DB 497 NIDVKNQAIIISVNDDFTA-AQLVW---PYTGPADKETPELMHVLRAQE-----AQIE 545

QY 1082 KVPQTD---PAFEVFPYTIYNVFEQVLTILPEGLFMLSICLPYTFVAVSCLILG-----L 1133
DB 546 DVTGTGLTGFTAVQLDIT-----EQLEDAMP-----VYLAVVVGLAIFLLILVFRSLLV 596

QY 1134 DLRSGLNLLSI-----VMILVDTVGFMAWD-----ISYNAVSLINLVSAGVMSV----- 1179
DB 597 PLVAGLGLLSVGAAGATVLVWQEGGGFVNTPGPLISFMPFLIGVTGTLANDYQVFL 656

QY 1180 -----EFVSHITRSFAISTKPTWLERAKEATISMGs-----AVFAGVAMTNLPG 1223
DB 657 VTRMRHVTHNGKQGPQSKVTPVEQSVIEGFTQGSRRVWTAALIMITAVFVAFDQPLPF 716

QY 1224 ILVGL--LAKAQLTQIFPRNLNLTILGLL-HGLVFLPVLTLSVVGDDVN-PALALQKR 1279
DB 717 IKIFGFALGAGVFFDAFFIRMLGPASFLMGKATWMPKWLDRILPSLDIEGTALKEW 776
```

```
QY 1280 ABEAVA 1285
DB 777 BEKQAA 782
```

RESULT 31

```
US-11-087-099-11792
; Sequence 11792, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 11792
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-11792
```

Query Match 1.5%; Score 105; DB 7; Length 324;

Best Local Similarity 23.2%; Pred. No. 1.5; Indels 136; Gaps 21; Matches 90; Conservative 48; Mismatches 114;

```
QY 527 NAPTFKDGKTALALSCMADYG-----APVPPFLAIGGYKGYSEAEALIMTFLNNY 579
DB 22 NGYLTGTGLFILLSLINVYGIKESARLAIPISFIEVSG-----LLIVIVT----- 67

QY 580 PAGDPRLAQAKLWEAFLEEMRAFORRMAGMEQVT---FTAERSLED-----EINR 627
DB 68 --GLPYLGTVNVYFETPDL-----AGIFKASTLIPFAFLGPDIVRLSQETKEAEK 115

QY 628 TTAEDLPFATSVIVFLYISLALGSYSWSRVMVDSKATLGLG-----VAVVLG--- 678
DB 116 TTPKALLI--AIFFTVFLMVCVATVVS-----MLDPR-VLGLSGAPLADVAISLGNKA 167

QY 679 -AVMAAMGPFYSYLGIRSSILVILQVVPFLVLSVADNIFIFVLEYORLPRRPGEPREVHIG 737
DB 168 FVLMGSHIALFSTMN-----TVLVMLGSSRIVYGMADSGALP----- 204

QY 738 RALGRVAPSMGLCSLSEALCFPLGALTMPAVRTFALTSGLA-----ILDPLLQ 787
DB 205 RFLSRVHP-----GYRTPWVALGVLSLFSALFVLFRDVAIVANISNFMIF 249

QY 788 MSAAFVALLSLDSKQREASRLDVCCCKVQPOELPPPGQEGLLGLFFQKAYAPFLLHMITRG 847
DB 250 IIFFWVNLSLIKLYTDPGRN-----RPFKVP-----LNIGRF--PLLPL-----G 289

QY 848 VVILLPLALFGVSL-----YSMCHISVGL 871
DB 290 ALUSTVFL-FQLSLEVITYGLVFIGIGI 316
```

RESULT 32

```
US-11-096-568A-27723
; Sequence 27723, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27723
; LENGTH: 1080
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
```

[illegible]

```

Qy 1099 NVFQEYL-----TI-----LPEGLF-----MLSLCLVFTFAVSC-LL 1130
Db 898 NVSRQVYVIGFICTIGASAGIGLVLSLIQLLFRKVFTKHTSSAVLDLANYQSLVATCVVL 957
Qy 1131 LGL-----DLRSLNLNLISVMILVDTVGFMAWLDI-SYNAVSLI-----N 1170
Db 958 IGLFASGEWRTLPEEMRNKYLKGVSYILTLASAAIF---WQVYTVGCVLIFESSVFSN 1014
Qy 1171 LVSAVGMSV 1179
Db 1015 SITAVGLPI 1023

RESULT 33
US-11-096-568A-27722
: Sequence 27722, Application US/110965568A
: Publication No. US20060048240A1
: GENERAL INFORMATION:
: APPLICANT: Alexandrov, Nikolai et al.
: TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
: TITLE OF INVENTION: Therby
: FILE REFERENCE: 2750-1592PUS2
: CURRENT APPLICATION NUMBER: US/11/096,568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471
: SEQ ID NO 27722
: LENGTH: 1097
: TYPE: PRT
: ORGANISM: Arabidopsis thaliana
: FEATURE:
: NAME/KEY: misc feature
: LOCATION: (1)..(1097)
: OTHER INFORMATION: Ceres Seq. ID no. 1825822
US-11-096-568A-27722

```

[illegible]

Db 456 RIQPKSTDTNFSQSPSTTLASVYLCTGLLVSAVAYLSALA-FTAFPSYFLNSQKFTPL 514
Qy 779 AVILDFLOMAFAVALLSLDSKQK-EASR-----LDVCCVKRQELPPPGQEGLLIGFPQ 833
Db 515 IVSSLLLTVTSALLVNTDSENSTNVRVQVIGICTI-----GASAGIGLLLSLIQ 568
Qy 834 KAYAPFLHWTIRGVVLLFLALFGVSLYSMCHISVGL-----DQELALPKDS 881
Db 569 MLFRKVTKHTSSAVT---DLAIY-QSLVASCVLIGLFGASGEWETLPSEMRNYKLGKVS 624
Qy 882 YLLD-----YFL-----FLNRYFEVGPVY----- 901
Db 625 YVLTLASAAISQVYTLGLVGLIFESSVFSNSITAVGLPIVPAVAIVFHRMDASKIP 684
Qy 902 -----FVTLTGVNFSSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db 685 SIILAIICGFLSFVYQHYLDEKCLNTSHTSAVGDHLHPVEEGHTNQSDQNLNLLDHEE 744
Qy 943 QSYLAIPASS-----WV-----DDFIDLTPSSCC--RLYISGPNKDKFCPTVNSLNCIK 991
Db 745 TESFSVPQTKNCKRMLRVSIYAI FVIFCQPLATVGLRLYENGSGKSTYVVTLLQIGFPV 804
Qy 992 NCWSTMTGMSVR--PSVEQPHKLPFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db 805 LILFRFSSRIQPKSTDTNFSQSPSTTLASVYLCT--GLLVSAVAYLSAVGLLYLPVST 862
Qy 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYIT 1098
Db 863 FSLILASQLAFTAFPSYFLNSQKFTPLIVNSLFLTVSSALL--VVNTDSE-----NTT 914
Qy 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 915 NVSRVQYVIGICTIGASAGIGLVLSLIQLLFRKVTKHTSSAVLDLANYQSLVATCVVL 974
Qy 1131 LGL-----DLRSGLLNLLSIYVILVDTVGFMAWDI--SYNAVSLI-----N 1170
Db 975 IGLFASGEWRTLPEMRNYKLGKVSILTASAAIF---WQVYTVGCVGLIFESSSVFSN 1031
Qy 1171 LVSAGMSV 1179
Db 1032 SITAVGLPI 1040
RESULT 34
US-11-096-568A-27721
; Sequence 27721, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 27721
; LENGTH: 1128
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1128)
; OTHER INFORMATION: Ceres Seq. ID no. 1825821
US-11-096-568A-27721

Query Match 1.5%; Score 105; DB 7; Length 1128;
Best Local Similarity 19.7%; Pred. No. 9.1;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

Qy 300 ILLVGFVPA-PARDKSKVDPKKGTSLSDKLSFSTHTLLGFOFGWG-----TWVASWP 352
Db 92 VQLIGFPVLFLFRFSSQTKNPK--PTBADFRKSSFTILGSVIVTGLLVSAVMSVSG 149

Qy 353 LTIIVLSVIPVVALAAGLVFTLTDTDPVELWSAPNSQARSEKAFDQHFQGPFFRTNQVIL 412
Db 150 LLXLPVSTFSLI-LASQLAFT-----AFFSYFLNSQKFTPIVNSLELL 192
Qy 413 TAPN-----RSSYRYSLLGLPKNFSGIL-----DLDLLELLEL-QEBRLRLQVWSP 459
Db 193 TISSALLVNTDSENTAKVSRVKVIGLICITIGASAGIGLLLSLVQLLIRKVLKQKTFE- 251
Qy 460 EAQRNLSLOD--ICVAPLNPNTSLYDCINSLLOYFQNN--RTLLLLTANQTLMGQTQSV 516
Db 252 ----TVTDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL-GKVPYV 296
Qy 517 -----DWKDHFLYCANAPLTFKQGTALALSCHADYGAPVFPFLAIGGVKKG-DYSEA 567
Db 297 MTLASIAISQVYTVGVG--LIFESSVFSNSITA-VGLFIVPVAIVFHDKNWASKI 353
Qy 568 EALIM-----TFSLNYY-----PAGDPRLAQAALWEBAFLERAFORRMAG 609
Db 354 FSIILAIWGFISFVYQHYLDEKCLNTSHTSPVGDPHLLPA---EEG----- 396
Qy 610 MFQVTFTAERSLE-----DEINRTAEDLPIFATSYIVIFLYISLALGSYSSWSRVMD 663
Db 397 --HTMIHSDQNLNLLIDHEVVTSESSAVP-----QTENYKRWLRVSIY 439
Qy 664 SKATILGLGVAIVLGAVMAMGFFS-YLGIRSSLVILQVVPFLVLSVGADNIFIFVLEYQ 722
Db 440 VIFVLCQPLATVILGRLYYENGSTYV-----VTLQLIGFPVLV---LFRP---PS 486
Qy 723 RLPRRGPPEVHIGRALGRVAPSMCLSLSEACIFFLGAITPMPAVRTEALTS-----GL 778
Db 487 RIQPKSTDTNFSQSPSTTLASVYLCTGLLVSAVAYLSALA-PTAFPSYFLNSQKFTPL 545
Qy 779 AVILDFLOMAFAVALLSLDSKQK-EASR-----LDVCCVKRQELPPPGQEGLLIGFPQ 833
Db 546 IVSSLLLTVTSALLVNTDSENSTNVRVQVIGICTI-----GASAGIGLLLSLIQ 599
Qy 834 KAYAPFLHWTIRGVVLLFLALFGVSLYSMCHISVGL-----DQELALPKDS 881
Db 600 MLFRKVTKHTSSAVT---DLAIY-QSLVASCVLIGLFGASGEWETLPSEMRNYKLGKVS 655
Qy 882 YLLD-----YFL-----FLNRYFEVGPVY----- 901
Db 656 YVLTLASAAISQVYTLGLVGLIFESSVFSNSITAVGLPIVPAVAIVFHRMDASKIP 715
Qy 902 -----FVTLTGVNFSSEAGMNAICSSA-----GCNPFSTQKIQ-YATEPPE 942
Db 716 SIILAIICGFLSFVYQHYLDEKCLNTSHTSAVGDHLHPVEEGHTNQSDQNLNLLDHEE 775
Qy 943 QSYLAIPASS-----WV-----DDFIDLTPSSCC--RLYISGPNKDKFCPTVNSLNCIK 991
Db 776 TESFSVPQTKNCKRMLRVSIYAI FVIFCQPLATVGLRLYENGSGKSTYVVTLLQIGFPV 835
Qy 992 NCWSTMTGMSVR--PSVEQPHKLPFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db 836 LILFRFSSRIQPKSTDTNFSQSPSTTLASVYLCT--GLLVSAVAYLSAVGLLYLPVST 893
Qy 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYIT 1098
Db 894 FSLILASQLAFTAFPSYFLNSQKFTPLIVNSLFLTVSSALL--VVNTDSE-----NTT 945
Qy 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 946 NVSRVQYVIGICTIGASAGIGLVLSLIQLLFRKVTKHTSSAVLDLANYQSLVATCVVL 1005
Qy 1131 LGL-----DLRSGLLNLLSIYVILVDTVGFMAWDI--SYNAVSLI-----N 1170
Db 1006 IGLFASGEWRTLPEMRNYKLGKVSILTASAAIF---WQVYTVGCVGLIFESSSVFSN 1062
Qy 1171 LVSAGMSV 1179
Db 1063 SITAVGLPI 1071


```
QY 242 RCNESQDDVATCSCODCAASCPAIPARQALDSTFVLGMPGSLVLIILCSVFAVVTIL 301
Db 1415 RYDLIVGE-----FCAEHGYDFESVNTAA-----SLAVLPFAENYFPLDWA 1457
QY 302 LVGFVAPARDKSKMVDPKGTSLSKLSFSHTLLGQFFQGWGTWVASWPLTIL-VLSV 360
Db 1458 AVYLLALLCRESNITSAADAGSGSPN-----GGIITGWISQVREVDEDLPLAA 1508
QY 361 IPVVALAAGLVTELTTPDVELMSAPNSQARSEKAFHQHFPFPRTRNQVILTAENRSY 420
Db 1509 LNVFQCAPG-DYTNVIGKTIE-----SGEFLLEDRLNALSWISGL 1548
QY 421 RYDLSLLGPKNFGSGLDLDLLELELOERLRLQWSPEAQRNISLQDICVAPLPDNT 480
Db 1549 SY-----VYGPKNHGASFPLLALN-LAAPDRTLHTWVSDEKA-----T 1587
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540
Db 1588 FFYDDCIYA-----FEGGLELAVSAVNGLLPQETMID-----ALVL 1624
QY 541 SCWADYGAPVFPFLAIGGYKGYSEAEALIMTFSLNYPAGDPRIQAQKLMEEAFLEEM 600
Db 1625 NLNRN-----AALGSYVGGDYSQ-ARRLAEEITQLMAAANPNLAP-----WLRDLASQL 1671
QY 601 RAPQRMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIFLYISLALGSYSWSRV 660
Db 1672 -----SAGSYLAALLAS-NMTSLFSTVTRTALATGNETAARYSLMLPDPDTAWYDL 1724
QY 661 MVDSKATLGLGVAVVLGAV-NAAMGFFSYLGRSLVILQVVPFVLVLSVGADNIFIFVL 719
Db 1725 M-----RTVFNVPYVGLRFHGYSG---AVELLKRLGVILIR--GWSTLALGLL 1767
QY 720 BYQRLPRRPG-----EPREVIHGRALGRVAPSMMLCSLSEACFFLIGALTPMPAVRTFALT 775
Db 1768 GWEGIFRRTASTLVNRE-----WLARYSPEGL-----FSLA 1799
QY 776 SGLAVL-----DFLLQMSAFVALLSLDSKQEASRLDVCC-----VKPQE 817
Db 1800 VSLLVITAYDRAQHKLSSEVLRSREFVALVR-DVIVPELLRGLVCCCPGCVGNPVVQQR 1858
QY 818 LPPPGQEGLLGFFOKAVAPFLHMI 844
Db 1859 LEALSQYGLVFN-LREANAVFPAANYM 1884

RESULT 43
US-11-126-313-34
; Sequence 34, Application US/11126313
; Publication No. US20050288489A1
; GENERAL INFORMATION:
; APPLICANT: Hirsch, Joel
; TITLE OF INVENTION: VOLTAGE-DEPENDENT CALCIUM CHANNEL BETA SUBUNIT FUNCTIONAL CORE
; FILE REFERENCE: P-6758-US
; CURRENT APPLICATION NUMBER: US/11/126,313
; CURRENT FILING DATE: 2005-05-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 34
; LENGTH: 2312
; TYPE: PRT
; ORGANISM: homo sapiens
US-11-126-313-34

Query Match 1.5%; Score 103.5; DB 7; Length 2312;
Best Local Similarity 18.2%; Pred. No. 33;
Matches 226; Conservative 161; Mismatches 436; Indels 421; Gaps 59;

QY 125 CHNCTSPNQSLFNVTRVAQLGAGQLPDAVVAEAFVQHSFAE-----QSDSC 172
Db 887 CHGNCDPTQQ-----BAGGGE--AVVTFEDRAHRQSORRRHRRVRTEGKSSA 935
QY 173 SRVRVPA---AATLAVGT-----MCGVYGSALCNAQ-----RWLNFQDGTGNGLA 214
```

```
Db 936 SRSASQERSLDEAMPTEGSKDHELGNHGAKEPTIQEERAQDLRRTNSLMVSRGSLA 995
QY 215 -----PLDITFILLEPGQAV-----GSGIQPLNEGVAR-----CNESQGD--- 249
Db 996 GGLDEADTFLVLPHPELEVGHVLTVEQEPGSGSEQALLGNVQDGMGRVISOSEPLSCI 1055
QY 250 ----DVATCSQDCAASCPAIPARQALDSTFVLGMPGSLVLIILCSVFAVVTIL---- 301
Db 1056 TANTDKATTESTSVTVAIPDV--DPLVDST-----VVHISNKT 1092
QY 302 --LVGFVAPARDKSKMVDPKK-----GTSL---SDKLSFSFST-----HTLLG-Q 339
Db 1093 GEASPLKEAREDEEBEVEKKQKKEKRETKAMVPHSSMFISTTNPTRRACHYIVNLR 1152
QY 340 FQOQGTWVASWPLTILVLSVIVVALAAGLVFTLTTPDVELWAPNSQARSEKAFHQ 399
Db 1153 YFE-----MCILLVIAASSIALAA-----EDFV-----LTNSRKNVLRIFYD 1190
QY 400 HFGPFRTNOVI-----LTAPNRSRYRYSLLGPKNFGSGLDLDLLELL----- 445
Db 1191 VFTGVFTFEMVIMKIDQGLILQDGSYPR-----DLWNILDFVVVVGVAVPALANA 1241
QY 446 -----ELQERLRHLQWSPEAQRNISLQDICVAPLPDNTSLYDCCINSLLQYFQN 496
Db 1242 LGTNKGRDIKTIKSLRVLRLRP-----LKTI---KRLPKLKAVFDCVVTSLKNVF-- 1289
QY 497 NRTLLLTANQTLMGQTSQVDMKDHFLYCANAPL-TFKDGTALALSCWADYGAPVFPFLA 555
Db 1290 -NILIVYKLFMFIFAVIAVQVLFKGFYCTDSSKDKTEK-----C 1328
QY 556 IGGY -----KGKDYSEAE-----ALIMTFSLNYPAGDPRIQAQKLMEEAFL 597
Db 1329 IGNYVDHEKNMEVKGREKREHFEHYDNIIWALLTLFTVST-GEQPQVLQ----- 1378
QY 598 EEMRAPORMAGMFQVTTFAERSLEDEINRTTAEDLPFATSYIIVIF-----LYISLAL 651
Db 1379 -----HSDVTEEDRGPSSRNM---EMSIQFYVYVVFVFFFPFIFVALLI 1422
QY 652 GSYSSWSRVWDSKATLGLGVAVVLGAVVAVMAAMGFFSYLGRS-----SILVILQVVP 703
Db 1423 ITFQSQDKQMB-ECSLEKNERACIDFAI-SAKPLTRYNPQNRHTFYQYRWHFVVSPTSFE 1480
QY 704 FLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEACFFLIGAL 763
Db 1481 YTIMAMIALNTVLMKKYSAP-----CTYELALKYLNIAF 1516
QY 764 TPMPAVRTFALTSGLA-----ILDFFLQMSAFVALLSLDSKQEASR--- 806
Db 1517 T-----MVPSLECVLKVIAFGFLNYPDRDTWNI FDFITVIGSITEIILTD SKLVNTSGFNM 1571
QY 807 --LDVCCCVKPOELPPPGQEGLLGFFOKAVAPFLHMIWTRGVLLLFL-ALFGVSLYS 863
Db 1572 SFLKLFRAARLKLRLQGGYTIIRLLWTFVQSEK--ALPVCULLIAMFFIYALIGQVFG 1629
QY 864 MCHISVGLDQELALPKDSYLLDYF-----LFLNRYFEVGAPEVYFVTTL 906
Db 1630 ----NKLDEESHINHNFRSFFGSLMLLFRSATGEAMQEI MSLGKGECPDPTAPS 1685
QY 907 GYNFSSEAGNNA-----ICSSAGCNP--SFTQKIQYATEFPEQSYLAIPASSWV 954
Db 1686 GQNERERCOTDLAYVYVVSFIPFCSFLMLNLFVAVIMDNFEYLTR--DSSIL---GPHIL 1740
QY 955 DDFID-W--LTPSSCCCLY-----ISGP-NKDKFCPSVNSLNCNKMCSITMGS 1000
Db 1741 DEFVRVWAEYDRAACRIIHYTEMYEMLTMSPLPLGLGKCPKSV-----AYKRLVLMMPV 1796
QY 1001 VRPSVEQFHKLPLFWFLNDRPNIKCPKGG-----LAAY 1032
Db 1797 AEDMTVHTFSTLMALIR TALDIKAKGADROQLDSELOKETLAIWPHLSQKMLDLVPM 1856
QY 1033 STSVNLTSPQGVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDDPA--- 1089
```

```
Db 1857 PKASDLTVGKIYAMIMDYK-----QSKVKQROOLEBQKQAPQRMPSLQEI 1911
Qy 1090 --FEVPPYITNVF-----YEQYLITLPEGLFMLSCLVPT 1123
Db 1912 ANAKALPYLQODPVSGLSGRSGYPSPLSPQDIFOLA-CMDPT 1954

RESULT 44
US-11-087-099-3980
; Sequence 3980, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 3980
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Pseudomonas putida KT2440
US-11-087-099-3980

Query Match 1.5%; Score 103; DB 7; Length 463;
Best Local Similarity 20.0%; Pred. No. 3.7;
Matches 130; Conservative 66; Mismatches 183; Indels 270; Gaps 32;

Qy 203 LNFQDGTN---GLAPLDTF-----HLLBPQAVGSGIQPLNEGVARCNEQDDVATCS 255
Db 1 MNTVGSQDGLAQGFQPRHVMTLSIAGIIGAGLPVSG----- 37
Qy 256 QDCAASCAPAIAPQALDSTFYLGQMPGSLVLIILCSVFVAVVTILLGFRVAPARDKSK 315
Db 38 -----HATAAGPATIISYFVAGTL---VVLVWMLGEMAV----- 70
Qy 316 MVDPKGTSLSKLSSTHTLQGFQGWGTWVASWPLFILVSLVPPVALAAGLVFTL 375
Db 71 -AHPDTG-----SFTYA--DQAGRAGWTIGLWFWVLVPIEALAAAGHVLNAW 120
Qy 376 TTDPVELWAPNSQARSEKAFDQHFQPRFRTNQVILTAPNRSSYRSDSLILGPKNFSGI 435
Db 121 FPQ-VDSW-----IFALASVILLA----- 138
Qy 436 LLDLLELLELQERLRLHQVMSPEAQRNISLDICVAPLNPNTSLYDCCINSLLQYFQ 495
Db 139 -----GTNLP-----SVAKYGE 150
Qy 496 NNRTLLLLTANQTL-----NGQTSQVDWKDHFLYCANAPLTFKDGATALASCMADYG--AP 549
Db 151 FEFWFALIKVTAIGLIGLGFALMDW-----LPNREVSGLS-GLMAEYGGFAP 198
Qy 550 -----VPFPL-----AIGYKGYSE--AEAL-----IMTF-----SL 576
Db 199 KGMSAVVGAFITVMFSGITGEATVIAASESDPSRNAKATRSVIRISTFYILSIFVII 258
Qy 577 NNPVAGDPRLAQAKLWEEAFLEBMR---AFQRMAGMFQVTFPTAERSLEDEINRTAED 632
Db 259 SVVPWNDPQAVVGVSTQRA-LEIMNTPNAAFWVDLVVLVAVTSCMNSSI----- 306
Qy 633 LPFATSYIVIFLYISLAL-----GSYSSWSRVVDSKATLGLGGVAVVLGAVMAAGGFF 687
Db 307 -----YIASRMWFSLAKGDAPAFLNKTSKGVFPRAAVFG-----STLIGRAIAVLNYP 355
Qy 688 SYLGIRSSLVILQVDFVLVSGADNIFIP-VLEYORLPRRPGEPREVIHIGRALGRVAPS 746
Db 356 APRG-----VFEPFLASSGAIALLVMVIAISQLRMR-----RRERENTE 396
Qy 747 M-----LLCSLSAICFEL-GALTPW---PAVRT-FALTSGLAVIDEL 785
Db 397 LKFRMWLFPLYTWAVIIFAGALANVMYTPHRAEVSSTLGLAIVISFL 445
```

```
RESULT 45
US-11-087-099-2723
; Sequence 2723, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2723
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Bacillus anthracis str. Ames
US-11-087-099-2723

Query Match 1.5%; Score 102.5; DB 7; Length 475;
Best Local Similarity 19.7%; Pred. No. 4.2;
Matches 66; Conservative 43; Mismatches 101; Indels 125; Gaps 15;

Qy 643 IFLYISLALGSY-----SSWSRVVMDSKATLGL-----GG--VAVVLGA- 679
Db 78 IYSYARAGEEYIGFNSAWGYWLAGILGNVATIMLLFSTLGYFPFPGNSVASIVGAS 137
Qy 680 -VMAAMGFPSYLGIRSSLV-----ILQVVPFLVLSVGADNIFIFVLEYORLPRRPG 729
Db 138 LLLLTWLPILFLGIREASIMNVIAITIGLVPVILFIVVMVTAFRWDTFTQDFW-----G 191
Qy 730 EPREVHIGRALGRVAPSMLLCSLSEACPLGALTMPMPAVRTFALTS-----GLAV 780
Db 192 E-GTISVSISILQVKNMLV---TLWVFIG-----VEGAVLSGRAKNSRDVGKAT 238
Qy 781 ILDFLLOMSAFVALLSLDSKQREASRLDYCCVKPQELPPGQGE----- 825
Db 239 VLGLILVMSIYIILISVLMSGAMTRGELSV-----LETSPSHVLEHVVTGAVAIN 290
Qy 826 -----GLLGLGFF-----OKAVAPFLILHMITRQVWVLLLF- 853
Db 291 IGLVASLVGTLLGWPLLVSEISHVAGKGVFPKVFTKNKQTFPHMALISNGVAQVIFI 350
Qy 854 LALPGVLSYMSCHISVGLDQELALPKDSYLLDYFL 888
Db 351 IVLFSESTYQIMYF-----IASTSLVPLYL 376

RESULT 46
US-11-087-099-5717
; Sequence 5717, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5717
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Halobacterium sp. NRC-1
US-11-087-099-5717

Query Match 1.5%; Score 102; DB 7; Length 324;
Best Local Similarity 22.5%; Pred. No. 2.7;
Matches 81; Conservative 41; Mismatches 116; Indels 122; Gaps 16;

Qy 526 ANAPLTFKDGATLALSCMADYGVAPFPFLAIGYKGYSEABALIMTFSLNNYPAGDPR 585
Db 7 AAGAAPAGQAVVAASAGAPVQGPQVFLAVSG-----AL----- 41
Qy 586 LAQAKLWEEAFLEEMRAFORRMAGMFQVTFPTAERSLEDEINRTAEDLPFATSYIVIFL 645
```


RESULT 49
US-11-096-568A-2445
; Sequence 2445, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 2445
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(727)
; OTHER INFORMATION: Ceres Seq. ID no. 15219049
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (629)..(629)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-2445

Query Match 1.5%; Score 100.5; DB 7; Length 727;
Best Local Similarity 21.8%; Pred. No. 11;
Matches 91; Conservative 47; Mismatches 162; Indels 117; Gaps 17;
QY 46 SGLMTLSNVCLSNTPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVLEASLSI 105
Db 221 SASPVSDMVSNAPDIK-----PVSVMGTPPLRTGPPGGANVLLNLSQVRQVMS 275
QY 106 TKALLTRCPACDNFVNLHCHNTCS-----PNQSLFI-----NTRVAQLGA----- 147
Db 276 SAALAGASSVGQSAVAMHMSNMISTGMATSLPSPQTVFSTGQGITSMAGSGALMGSAQ 335
QY 148 -GOLPAVAYEAFYQHSFARQSYDSCSRVRVPAATLAV-----GTWCGVYGSGALCNAQR 201
Db 336 TGQSGPG-----NNAFSPQTTN-----VASNLGVSQPMQGNQSHSGAMMQGGI 381
QY 202 WLN-----FQDGTGNGLAPLDITFHLLEP--GQAVGSGIQPLNEGVARCNSQGDVVA 252
Db 382 SMQNMMGLGQGNVSSGTGGMPT-----PGVGQAQSGIQQLGG-----S 423
QY 253 TCSCDCAASCAPAIARPPALDSTF-----YLGQMPGSLVLIILCSFVAVVTILLVGF 305
Db 424 NSSAPNMQLSQPSGAMQTSQSKYKVVWEGNLSGQSGQGPVLI-----TRLEGY 472
QY 306 RVAPARD-----KSKMVDPKGTSLSDKLSF---STHTLLGQFFQGWGT 346
Db 473 RSASASDSLAANWPTMQIVRLISQDHMNNKQYVGKADFLVFRAMSGHGLQLODKKLC 532
QY 347 WVASWPLTILVSVIPVVALAAGLVTELTDPVELW----SAPNSQARSEKAFHQ 399
Db 533 AVIQLPSQTLILLSVSDKACRLIGMLF-----PGDMVVFQPIPNQOQQOQOQLHQ 583

RESULT 50
US-11-096-568A-2444
; Sequence 2444, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471

; SEQ ID NO 2444
; LENGTH: 836
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(836)
; OTHER INFORMATION: Ceres Seq. ID no. 15219048
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (738)..(738)
; OTHER INFORMATION: Xaa is any aa, unknown or other
US-11-096-568A-2444

Query Match 1.5%; Score 100.5; DB 7; Length 836;
Best Local Similarity 21.8%; Pred. No. 13;
Matches 91; Conservative 47; Mismatches 162; Indels 117; Gaps 17;
QY 46 SGLMTLSNVCLSNTPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVLEASLSI 105
Db 330 SASPVSDMVSNAPDIK-----PVSVMGTPPLRTGPPGGANVLLNLSQVRQVMS 384
QY 106 TKALLTRCPACDNFVNLHCHNTCS-----PNQSLFI-----NTRVAQLGA----- 147
Db 385 SAALAGASSVGQSAVAMHMSNMISTGMATSLPSPQTVFSTGQGITSMAGSGALMGSAQ 444
QY 148 -GOLPAVAYEAFYQHSFARQSYDSCSRVRVPAATLAV-----GTWCGVYGSGALCNAQR 201
Db 445 TGQSGPG-----NNAFSPQTTN-----VASNLGVSQPMQGNQSHSGAMMQGGI 490
QY 202 WLN-----FQDGTGNGLAPLDITFHLLEP--GQAVGSGIQPLNEGVARCNSQGDVVA 252
Db 491 SMQNMMGLGQGNVSSGTGGMPT-----PGVGQAQSGIQQLGG-----S 532
QY 253 TCSCDCAASCAPAIARPPALDSTF-----YLGQMPGSLVLIILCSFVAVVTILLVGF 305
Db 533 NSSAPNMQLSQPSGAMQTSQSKYKVVWEGNLSGQSGQGPVLI-----TRLEGY 581
QY 306 RVAPARD-----KSKMVDPKGTSLSDKLSF---STHTLLGQFFQGWGT 346
Db 582 RSASASDSLAANWPTMQIVRLISQDHMNNKQYVGKADFLVFRAMSGHGLQLODKKLC 641
QY 347 WVASWPLTILVSVIPVVALAAGLVTELTDPVELW----SAPNSQARSEKAFHQ 399
Db 642 AVIQLPSQTLILLSVSDKACRLIGMLF-----PGDMVVFQPIPNQOQQOQOQLHQ 692

RESULT 51
US-11-087-099-4795
; Sequence 4795, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4795
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Oryza sativa
US-11-087-099-4795

Query Match 1.4%; Score 100; DB 7; Length 279;
Best Local Similarity 24.3%; Pred. No. 3.1;
Matches 46; Conservative 31; Mismatches 68; Indels 44; Gaps 9;
QY 1118 LCLVPTFAVSCLL--LGLDLRGLNLLISVIMLVDTVTFGMALWDISYNAVSLINLSAV 1175
Db 25 LCIIIP---VLCILEILFDKVRYSRNTKLSILVLVVG--VAVCTVTDVSVNSKGLLAIAV 80
QY 1176 ---GMSVEFVSHITRSPAIS-----TKPT-----WLERAKEATISMGSA 1211

```
Db      81  WSTAQQHVHLLQKYSILGSFNLGHTAPQAASLLILGPPVDFPWLNRRTDFTNYTTI 140
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy     1212 VPAGVAM-----TNLPGILVLGLAKAQLQIIFFRNLNLLITLGLL-----HGL-----VF 1257
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     141  VTFIVLSCTIAVGTNLQFICIGRFTAVSFQVLGHMKTILVLTGLFFGKEGLNFHV 200
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy     1258 LPVILSYVG 1266
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     201  LGMMLAVIG 209
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 52
US-10-204-252-10
; Sequence 10, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 3389
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-10

Query Match      1.4%; Score 100; DB 6; Length 3389;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches 74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;

Qy      762  ALTPMP-AVRTFALTSGLAVID--FLLQMSAFVALLSLDSKQREASRLDVCCVKPQEL 818
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2135  ALSELPETLETLTLLTATVGGIFFLMSA-----RGIGKMTLGMCCIIT---- 2181
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      819  PPGQGEGLLGFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVGLDQELALP 878
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2182  -----ASILLW-----YAQIQPHWIAAASIILEFFLIV-----LLIPEPEKQRTP 2220
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      879  KDSYLLDYFLNRYFEVGAPVYFVTTLGYNPSSSEAGMNAICSSAGCANNFSFTQKIQ--- 935
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2221  QDNQL-----TYVVIAT-----LTVVAATWANEMGFLEKTKKDL 2254
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      936  ----YATEPPEQSYLAI---PASSWVDDFDLWLTSSCCRLYISGNKDKFCPSTVN-SL 987
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2255  GLGSIATQQPESNILDILRPASW-----TLVAVATFTVTPMLRHSIENSSVNVSL 2306
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      988  NCLNKMCSITMGSVRPSVEQFHXYLPWFLNDRPNIKCPKGGLAAYS--TSVNLTSQGV 1045
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2307  TATAQATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYQVNPITLTAALFLL 2355
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy     1046  ASRFMAYHKPLKNSQDYTEALRAARELANITA-----DLRKVGTDPDFE 1091
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2356  VAHYAIIIGPGLQAKTREAQKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2406
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 53
US-10-204-252-6
; Sequence 6, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
```

```
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-6

Query Match      1.4%; Score 100; DB 6; Length 3391;
Best Local Similarity 21.0%; Pred. No. 1.1e+02;
Matches 74; Conservative 49; Mismatches 127; Indels 102; Gaps 17;

Qy      762  ALTPMP-AVRTFALTSGLAVID--FLLQMSAFVALLSLDSKQREASRLDVCCVKPQEL 818
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2137  ALSELPETLETLTLLTATVGGIFFLMSA-----RGIGKMTLGMCCIIT---- 2183
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      819  PPGQGEGLLGFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVGLDQELALP 878
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2184  -----ASILLW-----YAQIQPHWIAAASIILEFFLIV-----LLIPEPEKQRTP 2222
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      879  KDSYLLDYFLNRYFEVGAPVYFVTTLGYNPSSSEAGMNAICSSAGCANNFSFTQKIQ--- 935
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2223  QDNQL-----TYVVIAT-----LTVVAATWANEMGFLEKTKKDL 2256
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      936  ----YATEPPEQSYLAI---PASSWVDDFDLWLTSSCCRLYISGNKDKFCPSTVN-SL 987
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2257  GLGSIATQQPESNILDILRPASW-----TLVAVATFTVTPMLRHSIENSSVNVSL 2308
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      988  NCLNKMCSITMGSVRPSVEQFHXYLPWFLNDRPNIKCPKGGLAAYS--TSVNLTSQGV 1045
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2309  TATAQATVLMGLGK-----GWPLS-KMDIGVPLLAIGCYQVNPITLTAALFLL 2357
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy     1046  ASRFMAYHKPLKNSQDYTEALRAARELANITA-----DLRKVGTDPDFE 1091
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     2358  VAHYAIIIGPGLQAKTREAQKRAAGIMKNPTVDGITVIDLDPIP-YDPKFE 2408
;      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

RESULT 54
US-10-204-252-8
; Sequence 8, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 3391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
```


ORGANISM: Type A PWD circovirus
US-10-718-264-3

ORGANISM: Type A PWD circovirus
US-10-718-264-3

ORGANISM: Type A PWD circovirus
US-10-718-264-3

Query Match 1.4%; Score 99.5; DB 6; Length 577;
Best Local Similarity 18.4%; Pred. No. 9.5;
Matches 69; Conservative 48; Mismatches 116; Indels 141; Gaps 18;

QY 865 CHI--SVGLDDELA--LPKDSYLL-----DYFLFLNRYFEVGPV-----YFV 903
DB 93 CHIEKAGTDQNNKEYCSKEGHILIECGAPRNGKRSGLSTAVTILETGSLLVVAEQPP 152
QY 904 TILGYNFSS-----EAGMNAICSSAGCANNFSFTQIKIQAATER--PEQSY 945
DB 153 VTVVRNFRGLAELLKVSQKQKEDWKTAVHVIIVGPGCG-----KQWARNFAERDYY 206
QY 946 LAIPASW-----VDDPIDWLPSSCCRLVISGNKDKFCPCSTVNSLNCNKM 994
DB 207 WKPSRNKNWGDVGHGEVWVLLDDFYGLPWDDLLRLC-----DRY-PLVET----- 251
QY 995 SITMGSVRSVQGHKYLWFLNDRPNIKCPKGLAAY-STSVNLTSD-----CQVLA 1046
DB 252 -----KGGTVPFILARSILITISNQAPQEWYSSTAV 280
QY 1047 SRFMAYHKPKNSQDYTEALRAARELANITADLRKVPDPAFEVPPYITNVFY---- 1102
DB 281 PAVEALYRITTLQFWKTAGEQSTEVEPG-----RFEAVDPPCALFPYKINVFFVITS 334
QY 1103 -----EYVLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLL--NLLSI 1145
DB 335 WFLFLFIRVQDKFSLYNSQPHIILGCGCILERIAQACV--LDIGVIMEPQLVSVF 391
QY 1146 VMLVVDTVGFMAW 1159
DB 392 I-IWLEPINCLVLM 404

RESULT 60
US-10-506-454-1056
; Sequence 1056, Application US/10506454
; Publication No. US20060068386A1
; GENERAL INFORMATION:
; APPLICANT: Mezhevaya, Alexi I
; APPLICANT: Polushin, Nikolai N
; APPLICANT: Shcherbinina, Olga V
; APPLICANT: Shakova, Vera V
; APPLICANT: Malykh, Andrei G
; APPLICANT: Kozvavkin, Sergei A
; TITLE OF INVENTION: The Complete Genome and Protein Sequences of the Hyperthermophile
; TITLE OF INVENTION: Methanopyrus kandleri AV19 and Monophyly of Archaeal Methanogens
; TITLE OF INVENTION: and Methods of Use Thereof
; FILE REFERENCE: FID001
; CURRENT APPLICATION NUMBER: US/10/506,454
; PRIOR FILING DATE: 2004-08-31
; PRIOR FILING DATE: 2003-03-04
; PRIOR FILING DATE: 2003-03-04
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1722
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1056
; LENGTH: 716
; TYPE: PRT
; ORGANISM: Methanopyrus kandleri
US-10-506-454-1056

Query Match 1.4%; Score 99; DB 6; Length 716;
Best Local Similarity 17.2%; Pred. No. 14;
Matches 84; Conservative 88; Mismatches 159; Indels 156; Gaps 19;

QY 846 RGVLVLLFLALFGVSLYSNCHISVGLDDELAIPKDSYLLDYFLFLNRYFEVGPVYFVTT 905
DB 4 RTAILUSFLATVAMYAGLHNLNVEVDQTKYLPD----- 37
QY 906 LGYNFSSEAGMNAICSSAGCANNFSFTQIKIY--ATEFPQSYL-----AIPASSWVDDFI 958

DB 38 ----RFESMKQWVVERELGTSTKTLIVIEADDDVTRKPVLDYMRRIEDRLRSKPYVEN-- 92
QY 959 DMLTSSCCRLVISG--PNKDKFCPCSTVNSLNCNKMSTMGSVRSVQFHKYLPWFLN 1017
DB 93 -----VRGAPDVLRESPPA-----AVTMPGLRLPMSEMER----- 125
QY 1018 DRPNIKCPKGLAAYSTSVNLTSDQVLAASRFMAVHKPLKNSQDYTEALRAARELANIT 1077
DB 126 -----SKEMFVSKDHKVAIRV-----GLKSDADYRK-----VV 154
QY 1078 ADLRKVPDPAFEVPPYITN---VPYEQYLTILPEGLFMLSCLVLPVTFVAVSCLL---- 1130
DB 155 PDVRSLSLRDKPKSVKPADVTGSPAINYDFYRGLKD---LVTVTALVSAVAALVYVDF 211
QY 1131 -----LGLDLRSGLLNLSIYVLMILVDTVGFMAWLDISVNAVSLNLSYSAVGMSEFV- 1182
DB 212 RRWAPVGLTI---ILSAVAVVLMGYWLGAPPY-----ATVLTVMUGVGDIYVI 263
QY 1183 ---SHITRSFAISTKP-----TWLSRAKEATISMG---SAPVAGVAMTNLPGILVLGLA 1230
DB 264 FTLTRFOEYDIKGRAGKGEAILTAVRAGRAVLITGLTASAGFAALALSEFRMVSBIG-- 321
QY 1231 KAQLQIFFRNLNLTILGLHGLVFLPVILSYVGPDPNPALALEQKRAEAAVAVMVA 1290
DB 322 -----LGIVAGILTAVALTLLVLPSSLQSIPIGRKSSEKKEBSMKVL 363
QY 1291 SCP-NHP 1296
DB 364 SIPVRHP 370

RESULT 61
US-10-204-252-14
; Sequence 14, Application US/10204252
; Publication No. US20060062803A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as represented by the
; APPLICANT: Secretary, Department of Health and Human Services, Centers for
; APPLICANT: Disease Control and Prevention
; TITLE OF INVENTION: AVIRULENT, IMMUNOGENIC
; TITLE OF INVENTION: FLAVIVIRUS CHIMERAS
; FILE REFERENCE: 14114.0334U2
; CURRENT APPLICATION NUMBER: US/10/204,252
; CURRENT FILING DATE: 2002-02-16
; PRIOR APPLICATION NUMBER: 60/182,829
; PRIOR FILING DATE: 2000-02-16
; NUMBER OF SEQ ID NOS: 148
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3391
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
; OTHER INFORMATION: synthetic construct
US-10-204-252-14

Query Match 1.4%; Score 98; DB 6; Length 3391;
Best Local Similarity 19.8%; Pred. No. 1.6e+02;
Matches 69; Conservative 52; Mismatches 132; Indels 96; Gaps 15;

QY 762 ALTPMPAVRTFALTSGLAVIDPLQLQMSAFVALLSDSKRQBSRLDVCCKVPQBLPPP 821
DB 2137 AUSELPE-----TLETLLLTLLATVTGTFILMSRGIGKMTLGMCCIT----- 2183
QY 822 GQGEGLLGGFFQKAYAPFLHMTIRGVVLLLFALFGVSLYSNCHISVGLDDELAIPKDS 881
DB 2184 ---ASILLW-----YAOIOPHIAASIIULEFLIV-----LLIPEPEKQRTPODN 2225
QY 882 YLLDVFLLNRYFEVGPVYFVTTLGYNFSSEAGMNAICSSAGCANNFSFTQIKI----- 935
DB 2226 QL-----TYVVIAT-----LTVVAATWANEMGFLEKTKDGLG 2259

Db 712 L-----WLKRHPNSRLKFIHWPVPFNGPSQ 736

RESULT 67

US-10-858-730-118
; Sequence 118, Application US/10858730
; Publication No. US2005025568A1
; GENERAL INFORMATION:
; APPLICANT: Bailey, Richard B.
; APPLICANT: Blomquist, Paul
; APPLICANT: Doten, Reed
; APPLICANT: Driggers, Edward M.
; APPLICANT: Madden, Kevin T.
; APPLICANT: O'Leary, Jessica
; APPLICANT: O'Toole, George
; APPLICANT: Trueheart, Joshua
; APPLICANT: Walbridge, Michael J.
; APPLICANT: Yorgey, Peter S.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR AMINO ACID
; TITLE OF INVENTION: PRODUCTION
; FILE REFERENCE: 14184-030001
; CURRENT APPLICATION NUMBER: US/10/858,730
; CURRENT FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/475,000
; PRIOR FILING DATE: 2003-05-30
; PRIOR APPLICATION NUMBER: US 60/551,860
; PRIOR FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 118
; LENGTH: 294
; TYPE: PR1
; ORGANISM: Lactobacillus plantarum
US-10-858-730-118

Query Match 1.4%; Score 96; DB 6; Length 294;

Best Local Similarity 22.5%; Pred. No. 6.9;
Matches 49; Conservative 42; Mismatches 83; Indels 44; Gaps 9;

QY 1085 GTDPAFEVPPYITINNVYEQYITILPEGLFMLSCLVPTFAVSCILLGL--DLRGLLNL 1142
DB 76 TATSGF-----TNTFYIQALKLIP--VAVAAVLMQAVWISTLLGAVIHRRPSRLQV 126
QY 1143 LSIIVMLVDTVGFEMALWDISYNAVSLINLVSAVGSVEFVS-----HITRFAISTK 1194
DB 127 VSIIVLIGTILTAAGLFPIIT-----QALSPWGLMFLAACSACTMQFTASLGNL 179
QY 1195 P---TWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFPRLNLLITLGL 1251
DB 180 PLSKTWL-----LCLGA--FILIAVSPQLVTPATPTATV-----GWGLIALFSM 224
QY 1252 LHGLVFLPVLISVXGPDVNPALALEQKRAEEAAVAVMV 1289
DB 225 VPELVMSYLFMPYLEGIGLIGLSLSLEPLASIVVAVFLL 262

RESULT 68

US-11-087-099-4490
; Sequence 4490, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4490
; LENGTH: 479
; TYPE: PR1
; ORGANISM: Escherichia coli K12
US-11-087-099-4490

Query Match 1.4%; Score 96; DB 7; Length 479;
Best Local Similarity 18.7%; Pred. No. 14;
Matches 112; Conservative 74; Mismatches 175; Indels 238; Gaps 26;

QY 721 YQRLPRRPGEPREVIHIGRALG-----RVAPSMILCSLSEACFFLGAITPMPAVRTF 772
DB 6 HDRKPYTAGEAAIMAINSPINIAAOPGKTRLRKSL---KLQVVMGLAYLTPMTVFDTF 62
QY 773 ALTSGLA---VILDFLLQMSA--FVALLSLDSKREQASRLDVCCCVKPOELPPPGQEGRL 827
DB 63 GIVSGISDGHVPASVYLLALAGVLFTAISYKLVQV-----FPEAGSA--- 104
QY 828 LLGFFQKAYAP---FLHWHITRGVVLFLFLALFGVSLYSMCHISVGLDQELALPKDSYLL 884
DB 105 -YTYAQKSINPHVGFVGM-----SSLL 126
QY 885 DYFLFLNRYFEVGPVPTVTLGYNFSSEAGMNAICSSAGCWNFSFTOKIQIYATFPPQS 944
DB 127 DY-LFL-----PM-----INVLLA-----KIYLSALFPE-- 149
QY 945 YLAIPASSWDDFDLWLTSPSCCRLYISGPNKDKFCPSTVNSLNLCKNC-----MSI 996
DB 150 ---VPPVWVVTFAILLTAAN-----LKSVMNVANFNTLFLVLVQLSI 188
QY 997 TMGSVRPSVEQFHK-----YLPWFILNDRPNIKCPKGLAAYSTSVNLSGQVGLASR 1048
DB 189 MVVIFLVVQGLHKGEGVGTWSLQPPFISENAHL-----IPIITGATIVCFS 235
QY 1049 PMAYHKPLKNSQDYTEALRAARELANITA-----DLRKPVTDBA 1089
DB 236 FLGFDVAVTTLSEETPDARVIPK-AIFLTAIVYGVGFIAASFPMQLFPDPSRFXDPDAA 294
QY 1090 FEVPYITINNVYEQYITILPE-----GLFMLSCLVPTFAVSCILLGLDLRSGLLNL 1143
DB 295 -----LPEIALYVGKGLFQSIFLCTTF-VNLTASGLASHASVSRLL 334
QY 1144 STV-----MILVDYTFGFMALWDISYNAVSLIN---LVSAVG 1176
DB 335 YVMGRDNPVPERVGVYHPKWRTPALNVIMGVIALSALFFDLVLTATALINFGALVAFTF 394
QY 1177 MSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGI-LVLGLAKAQL 1234
DB 395 VNLSVFNHFRRKGMN--KSWKDHFYLLMPLVGLATVGLVWVNLESTSLTGLVWASL 451
RESULT 69
US-11-087-099-10324
; Sequence 10324, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10324
; LENGTH: 479
; TYPE: PR1
; ORGANISM: Escherichia coli O157:H7
US-11-087-099-10324
Query Match 1.4%; Score 96; DB 7; Length 479;
Best Local Similarity 18.7%; Pred. No. 14;
Matches 112; Conservative 74; Mismatches 175; Indels 238; Gaps 26;
QY 721 YQRLPRRPGEPREVIHIGRALG-----RVAPSMILCSLSEACFFLGAITPMPAVRTF 772
DB 6 HDRKPYTAGEAAIMAINSPINIAAOPGKTRLRKSL---KLQVVMGLAYLTPMTVFDTF 62
QY 773 ALTSGLA---VILDFLLQMSA--FVALLSLDSKREQASRLDVCCCVKPOELPPPGQEGRL 827
DB 63 GIVSGISDGHVPASVYLLALAGVLFTAISYKLVQV-----FPEAGSA--- 104


```
Db 165 PELRQLQALLQRPHQNTQTGTRPCWGSTHPRKASONEERAPLREOQCNPLRQKLPBGA 224
Qy 493 -----YFQNRRLILLTANQTLMTQSTQVDWKDHFLYCANAPLTKDGTALALS 541
Db 225 EAGTVLAGELGFLAOPGAFRLRPNVVLGSLTEVLSRFFCLLGP-----272
Qy 542 CWADYGAPVFPPLAIGGYKGDYS--AEALIMTFSLNNYPAGDPRLAQAKLWE-----593
Db 273 CML-----GKGHEMGRAAVLLS-----DPQFQ-----NSVRRAS 303
Qy 594 -----EAFLEEMRAP-----QRMAGMFOVTTFA 617
Db 304 NLHLLAALDALEEVTVLPGRWDPTARIPPKCLPSQHKELPSQREIRGPAVRLT- 362
Qy 618 ERSLED-----EINRT-----TAEDLPIFATSYI-----VIFLY 646
Db 363 --SAEDRRHGHGPHASPELQRTGRGLFQGLIQDVRKVPWYPSDFDLALHQCFSVLYI 420
Qy 647 ISLAGSYSSWSRVVDSKATILGLGV-AVVLGAVMAAGRFSYLG---IRSSLVILQV 701
Db 421 LATVTNAI--TFGGLGD--ATDGAOGVLESFUGTAVAGAACLMAGQPLTILSSGTVLV 477
Qy 702 VPFLVLSVGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSKMLLCSLSEALCFELG 761
Db 478 FERLLFSFSD---YSLDY--LPFR-----LWVGIIWA-----TFCLVLV 512
Qy 762 ALTPMPAVRTFA--LTSGLAVILDFLLQMSAFVALLSLDSK---ROEASRLDVCCKVRQ 816
Db 513 ATEASVLVRYFRFTBEGFCALISIFIYDAVGKMLNLTHTYPIQKPGSSAYGCLC----568
Qy 817 ELPPPGQGGGLLGPFOKAYAPFLHMI-----TGVVLLFLALFGVLSYMS-----864
Db 569 QYPPGPGNES-----QWTRPKDRDDIVSMDLGLINASLLPPPECTROG 613
Qy 865 -----CHISVGLDQELALPKDSYLLDYFLNRYFE-----VGAPYFVFTILGYNFS 911
Db 614 GHPRGPGCH-----TVP-DIAFFSLLFLTSPFAMALKCVKTSRFPSPVVRKGLS 663
Qy 912 SEAGMNAICSSAGCNF--SFTQKIQYATEP-----PEQSYLAIP--ASSWVDVDFDILTP 963
Db 664 DFFSVLAILLGCGDLAFLGLATPKLMVPREFKPTLFGRGLVSPFGANPMW-----WSVA 718
Qy 964 SSCRLYIS-----GPNKDFCPTSTVNSLNCMLKNCMSITM 998
Db 719 AALPALLSLIPMDQOITAVILNRMEYRLQKAGFHLDFCVAVLMLL-----TSAL 771
Qy 999 GSVRPSVEOFHKYLPFLWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKN 1058
Db 772 G-----LPWVVS-----ATVISLA-----785
Qy 1059 SQDYTEALRAARELANITADLRKVGTDPAP-----EVPFYITNVFYEQYLT 1107
Db 786 ---HMDSLRRESRACA-----PGERPNFLGIREQRLTGLVWFLTGASIFLAPVLK 833
Qy 1108 ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSI 1145
Db 834 FIP-----MP-----VLYGIFLYMGVAALSSI 855

RESULT 72
US-11-097-125-2
; Sequence 2, Application US/11097125
; Publication No. US2005026600A1
; GENERAL INFORMATION:
; APPLICANT: HADLEY, GREGG ALLEN
; TITLE OF INVENTION: THERAPEUTIC BLOCKADE OF CD103 INTERACTIONS TO PREVENT
; TITLE OF INVENTION: CLINICAL RENAL ALLOGRAFT REJECTION
; FILE REFERENCE: 0402.003.0002
; CURRENT APPLICATION NUMBER: US/11/097,125
; PRIOR APPLICATION NUMBER: 2005-04-04
; PRIOR FILING DATE: 2004-04-02
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 3.3
; SEQ ID NO 2
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Mus sp.
US-11-097-125-2

Query Match
1.4%; Score 96; DB 7; Length 1167;
Best Local Similarity 22.7%; Pred. No. 49;
Matches 96; Conservative 49; Mismatches 157; Indels 120; Gaps 19;

Qy 363 VVALAAGLVFTELTPVELMSPANSQARSKAFHQHGFPPFRTNQVILFAPNRSSTRY 422
Db 300 MVLVTGDIFF---GDPLNLTTVINSPKMQGVWFAIGVGRFK-----NNNTYRE 346
Qy 423 DSSLIG-PK-----NFSGILDLDLLELELEQLERLHLQ-----VMSPEAQRNLSL 467
Db 347 LKLIASDPKEAHTFKVTNYSY---LDGLLS--KLQQRIVHMEGTVDGLQVLAQTGFSA 401
Qy 468 QDICVAPLNPNTSLYDCCINSLLQYFONNRTLLL-----LTANQTLMTQTSQVDWK 519
Db 402 QILDKGVLLGTGAPNWSGGALLYSTQNGRCFLNQAKEDSRVQSYLGYSLAVLHK 461
Qy 520 DHFL-YCANAPLTFKOGTALALS-----CWADY---GAP 549
Db 462 AHGISYVAGAPRHKLGRGAVFELRKEDREDAFVRRIEQQMSGYFSGVLCVPVDIDMDGTT 521
Qy 550 VF-----PFLAIGGYKGDY-----SEAEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
Db 522 DFLVAAPFYHIRGEEGRVYVYQVPEQDA---SFLAHTLSGHGPGLTNSR-----568
Qy 601 RAFQRMAGMQVTFTAERSLEDEINRTTABDLPIFA-----TSYIVIFLYISLA 650
Db 569 --FGPMAAV-----GDLNQDKFTDVALGAPLEGFGAGDGASGYSVYIYNGHS 614
Qy 651 LGSYSSSRVMDVSKATILGLG--GVAVVLGAVMAAGRFSY-LGTRSSLVILQVVPFLVL 707
Db 615 GGLYDPSQOIRASSVAGSLHYFGMSVSGGLDFNGDLADITVGRSDRSLVSRPVDL 674
Qy 708 SV 709
Db 675 TV 676

RESULT 73
US-11-087-099-8709
; Sequence 8709, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 8709
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(338)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-8709

Query Match
1.4%; Score 95.5; DB 7; Length 338;
Best Local Similarity 21.7%; Pred. No. 9.2;
Matches 50; Conservative 41; Mismatches 84; Indels 55; Gaps 11;

Qy 1088 RAFEVFPYITITNVFYEQYLTILPEGLFML-----SLCLVPTFAVSCLL--LGLD 1134
Db 75 PSYLPFLVKKVFFPAN-LSIVGMNVSLMWSVGVGYQIAKLCII--XPVLCTLETWTWK 131
```

```

Qy 1135 LRSGLNLLSTVMILVDTVGPMALWDISYNASVLINLVSAV---GMSVEFVSHITRSPAI 1191
Db 132 VYRSRDTKLSITVLVNG--VAVCTVTDVSVNSKGLAAVIAVWSTALOOQYVHHLQKKYSL 190
Qy 1192 S-----TKPT-----WLERAKEATISMGSVAFAGVAM-----TNLPG 1223
Db 191 GSFNLLGHTAPAQAASLLILGPPVDFWLITNRRVDTFNTYITVTFVILSCTIAGVTNLQ 250
Qy 1224 ILVLGLAKAQLIQIFFRNLNLLITLGLL-----HGL---VFLPVLISYVG 1266
Db 251 FICIGRFTAVSFQVLGHMKITLVLTLGFLPFQKGLNFHVVLGMLMAVIG 300

RESULT 74
US-10-821-234-1027
; Sequence 1027, Application US/10821234
; Publication No. US20050255114A1
; GENERAL INFORMATION:
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Andarmani, Susan
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia
; FILE REFERENCE: 821A
; CURRENT APPLICATION NUMBER: US/10/821.234
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: US 60/462,047
; PRIOR FILING DATE: 2003-04-07
; NUMBER OF SEQ.ID NOS: 1704
; SOFTWARE: pt_seq_genes version 1.0
; SEQ ID NO 1027
; LENGTH: 686
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-821-234-1027

```

RESULT 75

US-11-096-568A-28283
; Sequence 28283, Application US/11096568A
; Publication No. US20060048240A1

```

: GENERAL INFORMATION:
: APPLICANT: Alexandrov, Nikolai et al.
: TITLE OF INVENTION: Sequence-Determined DN
: TITLE OF INVENTION: Theryb
: FILE REFERENCES: 2750-1592PUS2
: CURRENT APPLICATION NUMBER: US/11/096,568A
: CURRENT FILING DATE: 2005-04-01
: NUMBER OF SEQ ID NOS: 34471

```

```

; ORGANISM: Arabidopsis thaliana
; SEQ ID NO 28283
; LENGTH: 2186
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2186)

```

Query Match 1.4%: Score 95.5: DB 7: Length 2186:

Query Match 1.4%; SCORE 93.3; DB 7;
Best Local Similarity 19.7%; Pred. No. 1.3e+02;

Best local similarity	13.7%	Fied: NO. 1.3E+02;			
Matches	189;	Conservative	114;	Mismatches	305;
		Indels	351;	Gaps	48;

Qy	570	LIMTFSLNYPAGDRL-----AQAKLWBEAFUEMRAFORRMAGMFO	612
		::	::
Db	478	LVRQFSNNSLYDYIKLKALETWERINTKWAASAIWQSPSGATVTF-----AIYS	532
		::	::
Qy	613	VTFTAERSELEINRTTAEDLPFATSYVIVFLYISLALGSYSSWRVWVDSKATLGLGG	672
		::	::
Db	533	SILKSSQSYQS-----TLDFVSVPKDIGGTFGIISGLPYATMTSKSR-GEFG	579
		::	::
Qy	673	--VAVVLGAWMAAGFFSYLGRSSLVLQVVPFLVLSVGADNPIFVLEYQRLP----	725
		::	::
Db	580	PWVVVFLQVQVFGFFFIWASVGLTIAPPPVLMCL-----FVFLAGHS-LPFPNTA	631
		::	::
Qy	726	-----RRPGEPREVHIGRALGRVAPS--MLICSLSEATC-----FFLGALP----	765
		::	::
Db	632	NVVTARNPFGYGTAVGIMQGFGLSGAILQLYHAVCGEGNPATFILLAIIVPTLVN	691
		::	::
Qy	766	---MPAVRTFALTS-----GLA-----VILD	783
		::	::
Db	692	FLAMPFVRVYETVITSDKKHLDGLSAISMIYAALWVVITVENVLGSRSMQIPSFILVL	751
		::	::
Qy	784	FLIQMSFAVALLSLDSKROBASRLDVCCKVPQBLPPPG-----QGEGLLGFQKAYAP	838
		::	::
Db	752	LLLASPLLVAVRALREKROTILSSLDGPVLDTSSALLDPSSNIFPDGDHLV-----AEDS	805
		::	::

Qy	839	FLHWITRGVVLLFLAL	-----FGVSL-YSNCHISVGLDDELALPK- --	879
		: : : :		
Db	806	NILEAMSTNVFWLLFLAMLGMSGFATVNMNRQIGESLRYS	-----SVQLNSILWSLWSIW	861
		: : : :		
Qy	880	-----DSVLLDYFLNRYFEVGAPVPTVTLGYNFSSEAGMNAICSSAGCNPFSFTQ	932	
		: : : :		
Db	862	NFLGRFGAGVSDTFLKHSHW--PRPIFWAITLGV-----MAIGHIIVASGVQGSLYAG	913	
		: : : :		
Qy	933	KIQVATEFPBQ-----SYLAIAPASSWDDDFIDWLTPSSCCRLYISG	973	
		: : : :		
Db	914	SVLIGHMAYGSQMSLMPITITSEIFGIRHMGTYITISIAGPIGSYI	-----LSVKVIG	965
		: : : :		
Qy	974	PNKDKFCPTVNSL-----NCLKNCMSI-----TWGSVRPSVEQF--HKYLPWFMLNDRPNIK	1023	
		: : : :		
Db	966	YFYDKVASEDDNSCFGQCPRTSFMIMASVALFGSLVASVLPFRTHKFYKQLVAKR--NLK	1024	
		: : : :		
Qy	1024	CPKGGLAAYS-----TSVNLTSDCQVTLASRFMAYHKPLKN----	1058	
		: : : :		
Db	1025	SLDSLFSAAHAPPPVVVQPHIVLSLVFPKVSAGVLKRDG--LALRLVLNVLRLKSAETPE	1082	
		: : : :		
Qy	1059	-----SQDY-----TEALRAAEALANITAD--LRKVPGT	1086	
		: : : :		
Db	1083	CLISGLKCLVHLLTVBESIMWEGSDSYNILLNPFVTHSDGKVRKLASSCLRDVLQKSHGT	1142	
		: : : :		
Qy	1087	DPAFEPFPYITITNVFYEQYLITL-----PEG-----LFMLS-----LCUVPFVNS	1127	
		: : : :		
Db	1143	-KAWQSVSGAITEMF-QNYLDLAHKSEVSGTEGARGAKQVLYILSTLKECLALMSKKHIA	1200	
		: : : :		
Qy	1128	CLLLGLDLRSLGMLNLSIVMLIVDTVCFMALWDISYNVANSILINLVSAGVNSVEFVSHITR	1187	
		: : : :		
Db	1201	TLIBG-----FKVLMILRDPYITRVID--SLNAVCL-NPTSEV--PVEALDEV--	1244	
		: : : :		
Qy	1188	SFAISTKPTMLERAKEATISMGSAVPAG-----VAMTNLPGLIVLGLAKAQLIQIPFFRL	1242	
		: : : :		
Db	1245	-----LSUAGLFGSHETSADAMTFTARLLKVGWTRS-----FTL	1279	
		: : : :		
Qy	1243	N---LLITLGLLHGLFLPVLISYVGPDPNPALALEOKRAEAAVAV-----MVASC	1292	
		: : : :		
Db	1290	NPDICVVKIPASVENGI-----NDIITASEH-----EBAIPATDALKSLIPSC	1321	
		: : : :		

```

RESULT 76
US-11-096-568A-28282
; Sequence 28282, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nickolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28282
; LENGTH: 2199
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2199)
; OTHER INFORMATION: Ceres Seq. ID no. 2711347
US-11-096-568A-28282

```

	Query Match	1.4%; Score 95.5; DB 7; Length 2199;
	Best Local Similarity	19.7%; Pred. No. 1.3e-02;
	Matches 189; Conservative 114; Mismatches 305; Indels 351; Gaps	48;
Qy	570 LIMTSLNNYPAGDPRL-----AAAKLWEEAFLEENRAFQRWMGNFQ	612
Dd	LVQRFSSNLSYDYIKLKALKLETMERINTKWWAAAAIWIQSFGATYTF----IYS	545
Ov	613 VFTFAERSLEDENTTAEDLPFIATSIVIFLYISALGYSWSNRVVWDISKATLGUGG	672

RESULT 77

```

US-11-096-568A-28281
; Sequence 28281, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Pol
; TITLE OF INVENTION: Thery
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 28281

```


Db 18 GNMVSGIFSLPSSLA-----SIASPFQATSAWLL---TGAG 51
Qy 286 VLIITILCSFVAVVITILLNGFRVAP-ARDKSKWMDPKGTSLSKLSFSHTILLGQFFQW 344
Db 52 VLMIAL--VFCHLSIRKPELTAGPQSYARALFSDPKKGAAG-----FTWVW 96
Qy 345 GTWVASWPIITIVLSVIPVVALAAGLV--FTLETTDPVELMS 384
Db 97 GYVVASW-----ISNVAITSLAGYLTSFFPILVDKREMF 132
RESULT 82
US-11-087-099-10983
; Sequence 10983, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; PRIOR FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10983
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-10983

Query Match 1.4%; Score 95; DB 7; Length 552;
Best Local Similarity 17.4%; Pred. No. 20;
Matches 109; Conservative 82; Mismatches 185; Indels 250; Gaps 30;
Qy 635 IFATSVIVIFLYIS-----LALGSYSWSRWVDSKATLGLG-----GVAV 675
Db 52 VAGTMYIISLCASELSGATKAGPY-DWAKLFTGPGAAVSGLAVMEYALREADAI 110
Qy 676 VLGA-----VMAAMGFPSYLGIRSLV---ILQVVPFLVLSVGADNIF 715
Db 111 VVASISQSIFPELQVVPVTLVIALTFINRGVVAALTLNVLWTIAFIALA-----F 165
Qy 716 IFVLYQRLPRRPGPREVHIGRALGRVAPSMLLCSLSEAICFFIGALTPMPAVRTFALT 775
Db 166 PFSTAF-----GIGDIHPEYLL-----QCAL-PNGMIGLEA-- 195
Qy 776 SGLAVILDPLQMSAFVALLSDSKRQEARLDVCCVKPQELPPGQEGELGFFQKA 835
Db 196 -----ALQFGPWF-YLGIEGAMCAE-----CKHPSRAVPLGQOAGMIT----- 234
Qy 836 YAPFLLHMTIRGVVLLFLA-----LFGVSLYSNMCHISVGLDQELALPKDSYLLDYFL 888
Db 235 -----LLIGAAMTLYLCSVLIPADLLGVSVPL----- 262
Qy 889 FLNRPFEV---GAPVYVVTTLGN--FSSEAGMN-AICSSAG---CNNPFSFTQKIYATE 939
Db 263 -----FEAQNSGVFIYVALLGLTFLTCVASANGCVCDSSRLVLCSLKRLQRLVIRFSV 317
Qy 940 PPEQS--YLAI-----PASSWDDFDIWLTPSSCCRLYISGPNKDKFCPSTVNSL 987
Db 318 HPKNTYPRAVFTVPVPAIGFAPSGYLDQVITF-----SIVSGL 356
Qy 988 NC-----LKNCMSITMGSVRPSVEQPHKLPFWFLNDRPNIKCPKGGLAAYSTSVN 1037
Db 357 LCYVLIPFLIRFKLPETTSKVRFPVGPLOPIYAF-----AIAIA 399
Qy 1038 LTSDQGVILASRFWAYHKPL-----KNSODYTEALRA-----ARELA 1073
Db 400 IT-----ILSTLFWGYKYNLIFAFVSVLRTSISATNRTQTLKIGLKWAGPYPKAREIE 455
Qy 1074 ANITADLRKVPCTDPAFEPFPTIYNVFEQYL--TILPEGLFMLSCLVPTFAVSCILL 1131
Db 456 RSRRLQ-----ALETNDQALESKYHSLKTGDTALVVGMLALLICVEGFYAYKILSE 507

Qy 1132 GLDLRSGLNLLSI---VMILVDTVG 1154
Db 508 TWEIASGFFLYVIFIGLIVLIETIG 533
RESULT 83
US-10-505-263-8
; Sequence 8, Application US/10505263
; Publication No. US20060014940A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Case Western Reserve University
; APPLICANT: The Brigham and Women's Hospital, Inc.
; APPLICANT: Mount, David B
; APPLICANT: Romero, Michael
; TITLE OF INVENTION: CLONING AND CHARACTERIZATION OF SLC26A6, SLC26A1, and SLC26A2
; FILE REFERENCE: 1242/50/2 PCT/US
; CURRENT APPLICATION NUMBER: US/10/505,263
; PRIOR FILING DATE: 2004-08-20
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: PCT/US03/06469
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 735
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-505-263-8

Query Match 1.4%; Score 95; DB 6; Length 735;
Best Local Similarity 20.1%; Pred. No. 31;
Matches 124; Conservative 72; Mismatches 202; Indels 220; Gaps 32;
Qy 719 LEYQBLPRRPGPREVHIGRA-----LGRVAP-----SMLCSLSEAICFF-- 759
Db 1 MELQR-----RDYHVERPLNQLEDLGHGWPAAKTHQWRTWFRCSRARAHSLLLQ 52
Qy 760 -----LGALTPMPAVTFA---LTSGLAVILDPLQMSAFVALLSDSKRQEARLDVCC 812
Db 53 HVPVLGLPFRYP-VREWLLGLDLSGLSVAIMQLPQGLAYALLAGL----- 96
Qy 813 VKPQBLPPGQEGELGFFQKAYAPFLLHMTIRGVVLLFLALFGVSLYSMCHISVG-- 870
Db 97 -----PP-----MFLYSSFYVFIYP-----LFTSR-----HISVGF 126
Qy 871 -----LQDELALPKD-----SYLLDYFLFLNRYFEVGPVY---P 902
Db 127 AVMSVMVGSVTESLTADKAFVQGLNATADARVQVAYTILSPVLGL---FQVGLGLVHFGF 183
Qy 903 VTT-----LGVNFSSEAGMNAICSSAGCNFSFTQKIYATEPPEQSYLAIPASSWDDP 957
Db 184 VVTVLSBFLVRSYTTAAASVQVLVS-----QLKYVFGIKLSSHSG-PLS----- 225
Qy 958 IDWLTPSSCCRLYISGPNKDKFCPSTVNSLKNCMSITMGSVRPSVEQPHKLPFWFLN 1017
Db 226 VIYTVLEVCAQL-----PETVPGTVVTAIVAGVALVVLKLLNEKLHRLP---- 270
Qy 1018 DRPNIKCPKGGLAAYSTSVNLTSDQGVILASRFWAYHKPLKNSQDYTEALRAARELAANT 1077
Db 271 -----LPIGELLTLIGATGISYGVKLNDRPKV-----DVVGNIT 305
Qy 1078 ADLRKVPCTDPAFEPFPTIYNVFEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLRS 1137
Db 306 TGL--IPPVAPRTELEFATLVGNF-----AIAVVG-FAIAISLGKIFA---LHGYRVD 354
Qy 1138 -----GLNLLSIVMLVDTVGFMALWDISYN-AVSLINLVSAVMSVFEVSHITRSP 1189
Db 355 NOELVALGSLNIG-----GPFQCPVSCMSRSLVQ---ESTGGTQVAGAVSSLP 403
Qy 1190 AISTKPTWLERAKEATISMGSAVFAGVAMTNLPGL-----VLGLAKAQLIQIFPFLNL 1244


```
Db 483 LSLALISIMIGLICCCQOKT---GLPWWAPLLAVLLANILTIFFACWYGLVGY-----533
QY 653 SYSSWSRVMVDSKATLGLGGVAVVLGAVM-----AAMGFFSYLIGIRSSVLVQLVPPVL 707
Db 534 -YOPTSAVO-----MIGAYMVRPRPVANMMFTLYGSNALVQAIQMLGDLKL 578
QY 708 SVGADNIFVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALCFGLGALTMP 767
Db 579 A-----QYAKLP-----PRATFLAQILGTGAGSVFNMVWNSI-----611
QY 768 AVRTFALTSLAVILDFLLQMSAFVALLSDSKRQEARLDVCCVK-----814
Db 612 -----VDNR-----DILLSVQGTNWSGQNVQY 636
QY 815 -PQELPPPGQEGLL-----LGFQKAYAPFILHMITRGGVLLFLALFGVS 860
Db 637 NSQAVAWGGTANEILVHSGRHGTYMVMPLGFFGIPAP-LPFWI--GHKYFPQLRLNHN 693
QY 861 LYSMC-----HISVGLDQELALPKDSYLLDYFL-----LNRYFVGAPVYFVTTLYN 909
Db 694 TFLICTWLGWLSVGI-----NSSLLAYFVGFPSQGLRY-----KPALE-----734
QY 910 PSSEAGMNAICSSA---GCN-----NFSFTQKIYATEPEQSYLAIPASSW-----VDD 956
Db 735 -----AKWNNICAAIAGGCSLIIFILTFVAGGSGKARDPPK-----WMGNMMDG 780
QY 957 FIDMLTPSSCCRLYISG 973
Db 781 NVD-----RCKMYING 792
```

RESULT 86

```
US-11-087-099-4589
; Sequence 4589, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 4589
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Bacteroides thetaiotaomicron VPI-5482
US-11-087-099-4589
```

```
Query Match 1.4%; Score 94.5; DB 7; Length 231;
Best Local Similarity 21.9%; Pred. No. 6.4;
Matches 42; Conservative 35; Mismatches 70; Indels 45; Gaps 6;

QY 1106 LTLPLGLFMSLCVPTTAVSCLLGLDLRGLNLLSIVMLVDVTGFMALMDSYNA 1165
Db 1 MSFLENPFLLAV---TEGVFFAKLQKTKGLVLLNPILLITIAVLIIFLKMNISYET 56
QY 1166 VSLINLVSAGSVFVSHITRSPALSTKPTWLERAKEATISGSAVFAGVAMNLPGL 1225
Db 57 -----YNOGHHLIEF-----WL---RPAVALG-----VPLYL 81
QY 1226 VLGLAKAQIQLIFFFRNLNLLITLGLLHGLVFLPVLISYVGPDPVNPALALEOKRAEEAVA 1285
Db 82 QLEMIRKQLLP-----LLSQLAGCIVGVISVTLIAKFMGASQEVILSLAPKSVTTPTA 135
QY 1286 AVMVASCPNHS 1297
Db 136 MEVTKAIGGIPS 147
```

RESULT 87

```
US-11-087-099-2054
; Sequence 2054, Application US/11087099
```

```
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2054
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Methanosarcina barkeri
US-11-087-099-2054
```

```
Query Match 1.4%; Score 94.5; DB 7; Length 467;
Best Local Similarity 19.7%; Pred. No. 18;
Matches 78; Conservative 53; Mismatches 140; Indels 125; Gaps 19;

QY 635 IPATSYIVITFLYIS-----LALGSYSSWSRVMVDSKATLGLG-----GVAV 675
Db 52 VAGTWMYVILSLCASLHSGATKLGGPY-DWARLFIGPAAASVGLAVYMEYIALEADA 110
QY 676 VLGA-----VMAAMGPFESYLGRSSLV---ILQVVPFLVLSVGADNIF 715
Db 111 VVASISQSIPELQVYVPTLLVIALTLTFINRGVAAALTINFLVITWIAFIALA-----P 165
QY 716 IPVLEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSEALCFGLGALTMPAVRTFALT 775
Db 166 FESTA-----GQDHPHYLL-----QCAL-PNGMIGLFA-- 195
QY 776 SGLAVILDFLLQMSAFVALLSDSKRQEARLDVCCVKPQLPPPGQEG-----LLLGPF 832
Db 196 -----ALQFGPWF-YLGIEGAMCABE-----CKHPSRAVPLGQAGMITLLIGAA 240
QY 833 QKAY-----APFLHWTIRGTVLLFLALFGVLSYMSCHISVG---LDQE-- 874
Db 241 MTLXLCVLIIPADLLGVSVYPLFEAAQNSGIFIFVALLGLGTFLTCTVASANGCVCDSRS 300
QY 875 -LALPKDSYLLDYFLPLN-RYPEVGAPVYFVTTLGYNFSSEAGMN-----AICSSAGCNN 927
Db 301 WFALSRDNYVSSWFSANVHKYNTPTPRAVITPTVPAIGFAPSGVLDQVITFSIVGLLCYV 360
QY 928 FSFTQKIYATEFPQSYLAIPASSWVDDFIDWLT 963
Db 361 LIPPSLIRFKLFE-----TSKVRPFVGLQP 389
```

RESULT 88

```
US-10-479-873-9
; Sequence 9, Application US/10479873
; Publication No. US20060024673A1
; GENERAL INFORMATION:
; APPLICANT: EXELIXIS, INC.
; TITLE OF INVENTION: SLC2A8 AS MODIFIERS OF THE p53 PATHWAY AND METHODS OF USE
; FILE REFERENCE: EX02-076
; CURRENT APPLICATION NUMBER: US/10/479,873
; CURRENT FILING DATE: 2003-12-04
; PRIOR APPLICATION NUMBER: US 60/296,076
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/328,605
; PRIOR FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/357,253
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-479-873-9

Query Match 1.4%; Score 94.5; DB 6; Length 477;
Best Local Similarity 22.5%; Pred. No. 18;
```

Matches 108; Conservative 54; Mismatches 157; Indels 161; Gaps 27;
QY 526 ANAPUTFKDGTALASCADYCAPVFPFLAIGGYKDYSEAEALIMTFSL--NNY--PAGDPLAQAQKWEAA 595
Db 34 ALGPLSF--GPAFG--YSSPAIP-----SLQRAAPAPR 63
QY 586 LAQAKL--WEEAFLEEMRAFQORMAGMFOVTTAERSLEDEINRTTAEIDLPIFATSYIVIF 644
Db 64 LODAASWFGAVVTGGAAGGVIGW-----LVDRAGRLS--LLCSVPFFVAGF 111
QY 645 LYISIALGYSYSSWRSVMVDSKATLGLG--GVAVVLGAV--MAAMGFYSYLGIRSSLVILQV 702
Db 112 AVITAA---QDWM--MLLGGRLLTGLACGVASLAPVVISEIAYPAVRGLLGSCVQLMVV 166
QY 703 PFLVLSGADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA 762
Db 167 VGILLAYLAG---WLEWRML-----AVLGCVPPLMLL-----LMCF----- 201
QY 763 LTPMPAVRTEALTSLAVILDFLLQMSAFVALLSLDSKQEBASRLDVCCVKPQELPPPG 822
Db 202 ---MPETRFLLTQH-----RKQEAALRFLWSEQ-----GWEDDPFG 238
QY 823 QGEGLLGFFQK--AYAPFLHWTIRGVVLLFLALFGVSLY-----SM 864
Db 239 AEQSHLALLRQPGIYKPII-----GVSLMAFQQLSGVNAWMEYAEITFBEAKPKDSSL 293
QY 865 CHISVGLDQELAPKDSYLLD-----YFLFLNRYFEGAPVYVTT--LGYNFPSSBAGMNA 918
Db 294 ASWVGVIQVLTAVAAALIMDRAGRLLLVLS-----GVVMVFSTSAFGAYFKLTQG--- 345
QY 919 ICSSAGCNFTSQTKIYATEPEOSYLAI--PASSWDDF---IDWLTPSSCCRLYISG 973
Db 346 -----GPGN-----SSHVAISAPVSAQPVDSVGLAWLAVGSMC--LFIAG 384

RESULT 89

US-11-087-099-5936
; Sequence 5936, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5936
; LENGTH: 599
; TYPE: PRT
; ORGANISM: Ferroploasma acidarmanus
US-11-087-099-5936

Query Match 1.4%; Score 94.5; DB 7; Length 599;
Best Local Similarity 22.2%; Pred. No. 25;
Matches 125; Conservative 66; Mismatches 171; Indels 201; Gaps 33;

QY 283 GSULVLIILCSFPAVVTILLNFRVAPADKSKMDPKGTSLSDKLSSTHLLGQFPQ 342
Db 88 GAIVLVVL--VYAEIGAMI-----PRSG--FITRYGHYSHGGIAGLFP 127
QY 343 GNGTWTASWPLTILVLSVTPVVALAAGLVFT--ELTTDPVELMSAPNSQARSEKA----- 395
Db 128 GWGYFAAR-----VAAPALEAEAAITAGSYITKPAITYAKNPDPSSVTLISY 179
QY 396 -----PHDQHFGP--PFRTNQVI-----LTAPNRSYRYOSLLLGPKNFS----- 433
Db 180 GILIAAALTAGFYFLNFGVGLMGKTNQGITWKKLIIPSTIILMVFLLFHAGNFNNPAL 239
QY 434 -GIL-----DLDLLELLELQE-----RLRHLQWSPEA---QRNLSQDICTAPLNPDWT 480
Db 240 GGFLPHNNISLVEAISTDGVFSYLGFRHTLNFACEAKNPQDIP--RALIYAML-----T 294
QY 481 SLYDCCINSILQY--FQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTPKDGTA 539

Db 295 SI---VVVLLQAFIAINPSSLTASGWMGLSS-----ASA 329
QY 540 LSCWADYCAPVFPFLAIGGYKDYSEAEALIMTFSL--NNY--PAGDPLAQAQKWEAA 595
Db 330 GTYAKSIDSAFFAPLA-----KSSSLAILAVLYLLYDAYITSPAGTNLIAAGTATRSL 383
QY 596 F-LEEMRAFQORMAGMFOVTTAERSLEDEINRTTAEIDLPIFATSYIVIFLYISIALG-- 652
Db 384 YGLAEI-----GYFPRIF-----GVSKRTG--VPVFS-----LLISLVLSLI 419
QY 653 ---SYSSWSRVMVDSKATLGLGVAVVLGAVMAAMGFYSYLGIRSSLVILQVFPVLVSV 709
Db 420 FLVLPSPW-----YVVVGLVSGVAGTYILG--GSTLMVLR----- 453
QY 710 GADNIFIVLEYQRLPRRPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA-----LT 764
Db 454 -----REASDLKR-----PPKLPYARILSPVA-----FIGASLIVTWT 486
QY 765 PMPAVRTFA--LTSGLAVILDFL 785
Db 487 GWPTVAVIAIILFAGFAVLVFL 509

RESULT 90

US-11-205-109-4
; Sequence 4, Application US/11205109
; Publication No. US20050287641A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Zazopoulos, Emmanuel
; APPLICANT: Staffa, Alfredo
; TITLE OF INVENTION: GENE CLUSTER FOR RAMOPLANIN BIOSYNTHESIS
; FILE REFERENCE: 3002-ZUS
; CURRENT APPLICATION NUMBER: US/11/205,109
; CURRENT FILING DATE: 2005-08-17
; PRIOR APPLICATION NUMBER: US/09/976,059
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: US 60/239,924
; PRIOR FILING DATE: 2000-10-13
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Actinoplanes sp.
US-11-205-109-4

Query Match 1.4%; Score 94; DB 7; Length 336;
Best Local Similarity 23.1%; Pred. No. 12;
Matches 68; Conservative 44; Mismatches 126; Indels 56; Gaps 13;

QY 596 FLEEMRAFQORMAGMFOVTTAERSLEDEINRTTAEIDLPIFATSYIVIFLYISIALG----- 651
Db 65 YLAGVALVLPGLLGGFWGAPLITRELENGTQRLVWVNSVTRRRRWLLIKLLVVGLACKMVA 124
QY 652 ---GSYSKWSRVMVDSKATLGLGVAVVLGA---VMAAMGFYSYLGIRSSLVILQVVPF 704
Db 125 GVPSLLLWAAAPVDNVADNRF--TVMFGARFLPIAYAAFAVLGTILGLIVRETVA 182
QY 705 LVLSVGADNIFIVLEYQRLPR--RPGPREVHIGRALGRVAPSMMLCSLSEACFFLGA 762
Db 183 MALTIVAVFIQFL-----VPLNVRP-----HLPAPKHLVXP-MTVSAINEAKS--LGS 228
QY 763 LTPMPAVRTEALTSLAVILDFLLQMSAFVALLSLDSKQEBASRLDVCCVKPQELPPPG 822
Db 229 ITGAPVL-----NGLSISQGWITDVS---ALKTADGRSLDAKTFDNCY-----MNAKKTG 275
QY 823 QCEG-----LILGPFQKAYAPFLHWM-----ITRGVWLLFLALFGVSLY 862
Db 276 ATEPGYGDVAVCLAKLDLHDVIAYQPNRYWAFQFLESFYVLLSGLLGAAMV 329


```
Db 1340 ALAAVSYRLKQRTGFORQOKTLMSEWLL---ASSSPRISQTLCL-----MRP-QR 1387
Qy 606 RMAGMFQVTTFAERSLEDEINRT 628
Db 1388 RHPQSPPV--FPSQTLNRIWRT 1408

RESULT 93
US-11-082-389-16
; Sequence 16, Application US/11082389
; Publication No. US20050244935A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oekar
; APPLICANT: Haberer, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING PROTEINS
; TITLE OF INVENTION: INVOLVED IN MEMBRANE SYNTHESIS AND MEMBRANE
; TITLE OF INVENTION: TRANSPORT
; FILE REFERENCE: BGI-131CPGN
; CURRENT APPLICATION NUMBER: US/11/082,389
; CURRENT FILING DATE: 2005-03-16
; PRIOR APPLICATION NUMBER: US 09/603024
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: US 60/143262
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: US 60/151281
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930487.4
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19930489.0
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931549.3
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931550.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932134.5
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19941379.7
; PRIOR FILING DATE: 1999-08-31
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 446
; SEQ ID NO 16
; LENGTH: 296
; TYPE: PRP
; ORGANISM: Corynebacterium glutamicum
US-11-082-389-16

Query Match 1.4%; Score 93.5; DB 7; Length 296;
Best Local Similarity 21.6%; Pred. No. 11;
Matches 49; Conservative 45; Mismatches 76; Indels 57; Gaps 10;

Qy 1111 EGLFMLSCLVPTFAVSCLLLG-----LDLR-SGLNLN-----SIVMILVDTV 1153
Db 81 DGLLAGILTALWSNLNMGANVPLLRDNLFTPLRDAGLLGTWAGPAILAVAVGIL 140
Qy 1154 GFMALWDISYNAVSLNLINSAVCMSEFVSH---ITRSPAIKPTWLERAKEATISMG 1210
Db 141 GLVIWFLN-----TDGLSIRSTGDNQPMVQSGVSTDFT-----KILITISLN 185
Qy 1211 AVFAGVAMTNLP-----GILVLGLAKAQILQIFFRFLNLIITLLGLHLGLVFL 1258
Db 186 G-FVGLAGALIAQYQGFADISMIGLIVIGLASVILGQAFQRRVWLAVAVIUGAIY 244
Qy 1259 FVIL-----SVYQPDVNPALALEQKRAEEAAVAVMVASCPNHPRSVSTA 1302
Db 245 RLIIFAALRVGLDPN-----DMKAISSAILVVVAVMLLPRWAKFSKA 285
```

```
RESULT 94
US-11-087-099-2925
; Sequence 2925, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 2925
; LENGTH: 461
; TYPE: PRP
; ORGANISM: Shigella flexneri 2a str. 2457T
US-11-087-099-2925

Query Match 1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;

Qy 752 LSEAICFFLGALTMPAVRTALTSLGA---VILDFLLQMSA--FVALLSLDSKQOEASR 806
Db 24 LMQVVMGLAYLTMTVFTDFGIVSGISDGHVPASVYLLALAGVLFTAISYKLVQRQ---- 79
Qy 807 LDVCCCVKQELPPGQEGLLGFFQKAVAP---FLLHWITRGVWLLLLFLALFGVSLYS 863
Db 80 -----FPEAGSA---YTYAOKSINPHVGFVWGV----- 104
Qy 864 MCHISVGLDQELALPKDSYLLDYFLNRYPEVGAPVYFVTTGLGNFSSEAGMNAICSSA 923
Db 105 -----SSLDDY-LFL-----PM-----INVLLA-- 121
Qy 924 GCNNFSTQKQYATEPPEQSYLAIPASSWDDFIDMLTPSSCCRLYISGPNKDKFCPEST 983
Db 122 -----KIYLSALFPE-----VPPWVWVYVTFVAILTAAN----- 149
Qy 984 VNSLNLKNC-----MSITMGSVRPSVEQFHK-----YLPWFLNDRPNIKCPKG 1027
Db 150 LKSNVLVANFNTLVQVQISIMWVFIFLVVQGLHKGEVGVWSLOPFISENAHL----- 204
Qy 1028 GLAAYSTSVNLTSDQVLAIRFMAHKPLKNSQDYTEALRAARELANITA----- 1078
Db 205 -----IPIITGATIVCFSLGFDVAVTTLSEETPDAAVPEK-AIELTAVYGGVIFIA 255
Qy 1079 -----DLRKVPGTDPAFEVFPFYTITNVFEQYLTILPE-----GLFMLSCLVP 1122
Db 256 ASFFMQLFPPDISRFKPDAA-----LPEIALYVGGKLFQSFIFLCT 296
Qy 1123 TFAVSCLLGLDLRSLNLNLSIV-----MILVDTVGFMAL 1158
Db 297 TP-VNTLASGLASHASVSRLLYVMGRDNVFERVGVYHPKWRTPALNVIMVIGIALSAL 355
Qy 1159 WDISYNAVSLIN---LVSAGVMSVEFVSHITRSPAIKPTWLERAKEATISMGSAVFA 1215
Db 356 FFDLVATATALINFGALVAFTFVNLVSFNHFWRKGMN--KSWKDHFYLLMLPLVGLATVG 413
Qy 1216 VAMTNLPGI-LVLGLAKAQL 1234
Db 414 VLWVNLESTSLTGLVWASL 433

RESULT 95
US-11-087-099-4376
; Sequence 4376, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
```

```

; SEQ ID NO 4376
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Escherichia coli
US-11-087-099-43

```

Query Match 1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;

```

RESULT 96
US-11-087-099-9669
; Sequence 9669, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 9669
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Escherichia coli O157:H7 EDL933
US-11-087-099-9669

```

Db	24	LMQVVMGLAYLTPMTVDFGIVSGISDGHVPASVYLLALAGVLTFAISYGLVRQ----	79
Qy	807	LDVCCCVKQBELPPPGOGEGLLGFQKAYAP---FLHWHITRGVLLLLFLALFGVSLYS	863
Db	80	-----PPEAGSA---YTYAKSINPHVGFWMGV-----	104
Qy	864	MCHISVGLDQELALPKDSYLLDDYFLFLNRRYFEVGAPVYFVTTLGYNFSSEAGNAICSSA	923
Db	105	-----SSLDY-LFL-----PM-----INVLLA-----	121
Qy	924	GCNNFSTOKIQATEPEQSYLAIPASSWVDVDFIDWLTPSSCRLYISGPNKDKFCPST	983
Db	122	-----KIYLSALFPE-----VPPWVWVTFVAILTAAN-----	149
Qy	984	VNSLNCUKNC-----MSITGVSVRPSVEQFHK-----YLPWFLNDRPNKICPKG	1027
Db	150	LKSNNLVANFTLFLVQISIMVFIPLVVOGLHKGEVGVTWVLSLQPFISSENAHL-----	204
Qy	1028	GLAAYSTSUNLTSDDQVLSARFMAYHKPLKNSODYTTEALBAARELANITA-----	1078
Db	205	-----IPIITGATIVCFSLGFDVAVTTLSBETPDAARVIPK-AIFLTAVYGGVIFIA	255
Qy	1079	-----DLRKVGTDPAPFVPPPTITNVYEQYLTLPE-----GLFWLSLCLVP	1122
Db	256	ASPFMQLFPDISRFKDPDAA-----LPIBALYVGKLFQSFILCT	296
Qy	1123	TFAVSCILLGLDLRSGLLNLSLV-----MLVDVTGVFMAL	1158
Db	297	TF-VNTLASGLASHASVSRLLYMGDRNVPPERFVGYPHKWTPALNVMGVAMGAL	355
Qy	1159	WDISYNAVSLIN---LVSAVGMSVBFVSHITRSFAISTKPTWLERAKEATISMGSAVPAG	1215
Db	356	FEDLVATATALINFGALVAFTFVNLSVFNHFWRKGN--KSWKDHPHYLLMPLVGLATVG	413
Qy	1216	VAMTNLPGI-LVGLAKAQL	1234
Db	414	VLWVNLESTSLTGLVWASL	433

RESULT 97
US-11-087-099-10340
; Sequence 10340, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 10340
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(602)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-10340

Query Match 1.4%; Score 93.5; DB 7; Length 461;
Best Local Similarity 18.6%; Pred. No. 21;
Matches 104; Conservative 70; Mismatches 159; Indels 227; Gaps 24;
QY 752 LSAICFFLGALTPMPAVRTALTSGIA---VILDFLLOMSA--FVALLSDSKQREASR 806


```
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 24
; LENGTH: 4868
; TYPE: PRT
; ORGANISM: Meleagris gallopavo
US-11-044-111-24
```

```
Query Match      1.4%; Score 93.5; DB 7; Length 4868;
Best Local Similarity 20.2%; Pred. No. 5.9e+02;
Matches 80; Conservative 63; Mismatches 121; Indels 133; Gaps 25;

QY      893 YPEVGAPVVFVTTLGYNESSEAGMNAICSSAGCANNFSTQKIQYATEPPEQS-----YLA 947
Db      1286 YCLSLNPTEFRSS--FNFG--VGVENASSDA-----LQKRKHSQEFFPASSTTYFYSLR 1334
QY      948 I-----PASSWDDFDIDWLTTPSSCCRLYISGPNKDKFCPCSTVNSLNLCLKNC-MSITWGS- 1000
Db      1335 IPAGQDPSSVWV-----GWVTPD----YHFYSENF-----IN-----KNCIVTVTLGDE 1375
QY      1001 ---VRPVEQFHKYLPW-----FLNDRPNIKCPKGGLAAYSTS-VNLTSDGQVLAJR 1048
Db      1376 RGRVHESVRSNCYMWGCGDITANSQSRGSRNVLDLEIGCFVDLATGMLSTANGKELGTC 1435
QY      1049 EM-----AYHKP-----LKNQSDYTEALRAARELAANITADLRKVPCT 1086
Db      1436 YQVEPNTKLLPAFVQPTSTNLIQFELGKLKNTMPLSAAFKSEE-----RNEVPQC 1487
QY      1087 DPAFEVFPYTIITNVFYEQYLITLPEGLFMLS-----CLVPTFAVSCLLGLD 1134
Db      1488 PPRLDV--QTITPVLWSR----MPNSFLKVEVERSERHGWVQCLEP---LQMMALHIP 1538
QY      1135 LRSGLNLNLSIVMILVDTVGFMAWMDISYNAVSLINLVSAGV---MSVEFVSHITRS---- 1188
Db      1539 EENRCVDILE----LCEQEDLM---KPHYHTLKLYSSVCALGNTRVAYALCSHVDISQLF 1591
QY      1189 FAISTK-----PTWLERAKEATISMG 1210
Db      1592 YTIDNQYLPGLLRSGFYDLLISIHLDHAKQAKLMNN 1628
```

```
Search completed: April 11, 2006, 01:01:32
Job time : 53 secs
```

THIS PAGE IS BLANK

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2006, 00:46:22 ; Search time 556 Seconds
(without alignments)
3310.727 Million cell updates/sec

Title: US-10-736-769-4

Perfect score: 6909

Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKGAGAIENFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 7861189 seqs, 1381955077 residues

Total number of hits satisfying chosen parameters: 7861189

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :
1: /cgm2_6/ptodata/1/paa/US066 COMB.pcp:*
2: /cgm2_6/ptodata/1/paa/US073 COMB.pcp:*
3: /cgm2_6/ptodata/1/paa/US074 COMB.pcp:*
4: /cgm2_6/ptodata/1/paa/US075 COMB.pcp:*
5: /cgm2_6/ptodata/1/paa/US076 COMB.pcp:*
6: /cgm2_6/ptodata/1/paa/US077 COMB.pcp:*
7: /cgm2_6/ptodata/1/paa/US078 COMB.pcp:*
8: /cgm2_6/ptodata/1/paa/US079 COMB.pcp:*
9: /cgm2_6/ptodata/1/paa/US080 COMB.pcp:*
10: /cgm2_6/ptodata/1/paa/US081 COMB.pcp:*
11: /cgm2_6/ptodata/1/paa/US082 COMB.pcp:*
12: /cgm2_6/ptodata/1/paa/US083 COMB.pcp:*
13: /cgm2_6/ptodata/1/paa/US084 COMB.pcp:*
14: /cgm2_6/ptodata/1/paa/US085 COMB.pcp:*
15: /cgm2_6/ptodata/1/paa/US086 COMB.pcp:*
16: /cgm2_6/ptodata/1/paa/US087 COMB.pcp:*
17: /cgm2_6/ptodata/1/paa/US088 COMB.pcp:*
18: /cgm2_6/ptodata/1/paa/US089 COMB.pcp:*
19: /cgm2_6/ptodata/1/paa/US090 COMB.pcp:*
20: /cgm2_6/ptodata/1/paa/US091 COMB.pcp:*
21: /cgm2_6/ptodata/1/paa/US092 COMB.pcp:*
22: /cgm2_6/ptodata/1/paa/US093 COMB.pcp:*
23: /cgm2_6/ptodata/1/paa/US094 COMB.pcp:*
24: /cgm2_6/ptodata/1/paa/US095 COMB.pcp:*
25: /cgm2_6/ptodata/1/paa/US096 COMB.pcp:*
26: /cgm2_6/ptodata/1/paa/US097 COMB.pcp:*
27: /cgm2_6/ptodata/1/paa/US098 COMB.pcp:*
28: /cgm2_6/ptodata/1/paa/US099 COMB.pcp:*
29: /cgm2_6/ptodata/1/paa/US100 COMB.pcp:*
30: /cgm2_6/ptodata/1/paa/US101 COMB.pcp:*
31: /cgm2_6/ptodata/1/paa/US102 COMB.pcp:*
32: /cgm2_6/ptodata/1/paa/US103 COMB.pcp:*
33: /cgm2_6/ptodata/1/paa/US104 COMB.pcp:*
34: /cgm2_6/ptodata/1/paa/US105 COMB.pcp:*
35: /cgm2_6/ptodata/1/paa/US106 COMB.pcp:*
36: /cgm2_6/ptodata/1/paa/US107 COMB.pcp:*
37: /cgm2_6/ptodata/1/paa/US108 COMB.pcp:*
38: /cgm2_6/ptodata/1/paa/US109 COMB.pcp:*
39: /cgm2_6/ptodata/1/paa/US110 COMB.pcp:*
40: /cgm2_6/ptodata/1/paa/US111 COMB.pcp:*
41: /cgm2_6/ptodata/1/paa/US112 COMB.pcp:*
42: /cgm2_6/ptodata/1/paa/US113 COMB.pcp:*
43: /cgm2_6/ptodata/1/paa/US114 COMB.pcp:*

44: /cgm2_6/ptodata/1/paa/US600 COMB.pcp:*
45: /cgm2_6/ptodata/1/paa/US601 COMB.pcp:*
46: /cgm2_6/ptodata/1/paa/US602 COMB.pcp:*
47: /cgm2_6/ptodata/1/paa/US603 COMB.pcp:*
48: /cgm2_6/ptodata/1/paa/US604 COMB.pcp:*
49: /cgm2_6/ptodata/1/paa/US605 COMB.pcp:*
50: /cgm2_6/ptodata/1/paa/US606 COMB.pcp:*
51: /cgm2_6/ptodata/1/paa/US607 COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6909	100.0	1332	1	PCT-US03-22467-4
2	6909	100.0	1332	1	PCT-US03-40113-4
3	6909	100.0	1332	1	PCT-US05-01469-4
4	6909	100.0	1332	36	US-10-621-758A-4
5	6909	100.0	1332	36	US-10-646-301A-4
6	6909	100.0	1332	36	US-10-663-208A-4
7	6909	100.0	1332	37	US-10-736-769-4
8	6909	100.0	1332	37	US-10-750-386-4
9	6896	99.8	1332	1	PCT-US01-04098A-1831
10	6896	99.8	1332	32	US-10-239-316-9
11	6896	99.8	1332	32	US-10-258-899A-1831
12	6896	99.8	1332	32	US-10-293-244-1831
13	6896	99.8	1332	42	US-11-218-141-1831
14	6896	99.8	1332	42	US-11-242-459-9
15	6872.5	99.5	1359	1	PCT-US01-04098A-1830
16	6872.5	99.5	1359	1	PCT-US03-22467-44
17	6872.5	99.5	1359	1	PCT-US03-40113-44
18	6872.5	99.5	1359	1	PCT-US05-01469-44
19	6872.5	99.5	1359	1	PCT-US05-27579-22
20	6872.5	99.5	1359	32	US-10-258-899A-1830
21	6872.5	99.5	1359	32	US-10-293-244-1830
22	6872.5	99.5	1359	36	US-10-621-758A-44
23	6872.5	99.5	1359	36	US-10-646-301A-44
24	6872.5	99.5	1359	36	US-10-663-208A-44
25	6872.5	99.5	1359	37	US-10-736-769-44
26	6872.5	99.5	1359	37	US-10-750-386-44
27	6872.5	99.5	1359	42	US-11-218-141-1830
28	6868.5	99.4	1359	31	US-10-170-205E-21686
29	6868.5	99.4	1359	48	US-60-453-050-8956
30	6868.5	99.4	1359	48	US-60-466-412-8956
31	6868.5	99.4	1359	48	US-60-466-412-8956
32	6536	94.6	1344	1	PCT-US01-08631-53052
33	6536	94.6	1344	34	US-10-450-763-53052
34	5421.5	78.5	1331	1	PCT-US03-22467-2
35	5421.5	78.5	1331	1	PCT-US03-40113-2
36	5421.5	78.5	1331	1	PCT-US05-01469-2
37	5421.5	78.5	1331	36	US-10-621-758A-2
38	5421.5	78.5	1331	36	US-10-646-301A-2
39	5421.5	78.5	1331	36	US-10-663-208A-2
40	5421.5	78.5	1331	37	US-10-736-769-2
41	5421.5	78.5	1331	37	US-10-750-386-2
42	5407	78.3	1333	1	PCT-US03-22467-12
43	5407	78.3	1333	1	PCT-US03-40113-12
44	5407	78.3	1333	1	PCT-US05-01469-12
45	5407	78.3	1333	1	PCT-US05-27579-3
46	5407	78.3	1333	36	US-10-621-758A-12
47	5407	78.3	1333	36	US-10-646-301A-12
48	5407	78.3	1333	36	US-10-663-208A-12
49	5407	78.3	1333	37	US-10-736-769-12
50	5407	78.3	1333	37	US-10-750-386-12
51	4466	64.6	982	1	PCT-US01-08631-53050
52	4466	64.6	982	34	US-10-450-763-53050
53	2403	34.8	1318	39	US-10-940-774-10152
54	2403	34.8	1318	39	US-10-940-774A-10152
55	2402.5	34.8	1278	1	PCT-US03-40977-530


```
; FILE REFERENCE: 36134-PCT 074669, 0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US 60/537,341
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01469-4

Query Match      100.0%; Score 6909; DB 1; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db      1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60

Qy      61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db      61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120

Qy      121 VNLCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180
Db      121 VNLCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180

Qy      181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db      181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240

Qy      241 ARCNEQGDVATSCDQCAACPAIRPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db      241 ARCNEQGDVATSCDQCAACPAIRPAQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

Qy      301 LLVGRVAPAROKSRVDPKGTSLSDKLSFSTHLLGQFFQGWGTWASWPLTILVLSV 360
Db      301 LLVGRVAPAROKSRVDPKGTSLSDKLSFSTHLLGQFFQGWGTWASWPLTILVLSV 360

Qy      361 IPWVLAAGLVTELTDDPVELWSAPNSQARSEKAFDHOHGFPPFRTNQVILTAENRSY 420
Db      361 IPWVLAAGLVTELTDDPVELWSAPNSQARSEKAFDHOHGFPPFRTNQVILTAENRSY 420

Qy      421 RYDSSLGPKNFGSLDLDLLELELQRLRLQVWSPEAQRNISLQDICVAPLNPDNT 480
Db      421 RYDSSLGPKNFGSLDLDLLELELQRLRLQVWSPEAQRNISLQDICVAPLNPDNT 480

Qy      481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTFCQGTALAL 540
Db      481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFCANAPLTFCQGTALAL 540

Qy      541 SCWADYGAPVFPFLAIGGKGDYSABEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
Db      541 SCWADYGAPVFPFLAIGGKGDYSABEALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600

Qy      601 RAFQRMAGMFWQVTFRTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
Db      601 RAFQRMAGMFWQVTFRTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660

Qy      661 MVDSKATLGGVAVVILGAVMAAMGFFSYLIGIRSSLVILQVPPFLVLSVADNIFIFVLE 720
Db      661 MVDSKATLGGVAVVILGAVMAAMGFFSYLIGIRSSLVILQVPPFLVLSVADNIFIFVLE 720

Qy      721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSSEALCFPLGALTMPAVRTFALTSGLAV 780
Db      721 YQRLPRRPGEPREVHIGRALGRVAPSMMLCSSEALCFPLGALTMPAVRTFALTSGLAV 780

Qy      781 ILDFLLQMSAFVALLSLDSKRBASRLDVCCCKVQPELPPPGQEGILLIGFFQKAYAPFL 840
Db      781 ILDFLLQMSAFVALLSLDSKRBASRLDVCCCKVQPELPPPGQEGILLIGFFQKAYAPFL 840
```

```
Qy      841 LHMITRGWVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db      841 LHMITRGWVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

Qy      901 YFVTTLGYNFSSEAGMNAICSSAGCNPNFPTQKIQYATEPPEQSYLAIPASSWDDDFIDW 960
Db      901 YFVTTLGYNFSSEAGMNAICSSAGCNPNFPTQKIQYATEPPEQSYLAIPASSWDDDFIDW 960

Qy      961 LTPSSCCRLYISGPNKDKPCPSTVNSLNCNKNCMSITMGSVRPSVEQPHKYLPLNDRP 1020
Db      961 LTPSSCCRLYISGPNKDKPCPSTVNSLNCNKNCMSITMGSVRPSVEQPHKYLPLNDRP 1020

Qy      1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db      1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080

Qy      1081 RKVPTDPAPEVPFPTITNVFEQYLTILPGLPMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db      1081 RKVPTDPAPEVPFPTITNVFEQYLTILPGLPMLSCLVPTFAVSCLLGLDLRSGLL 1140

Qy      1141 NLLSIVMLTVDVTFGMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKETWLER 1200
Db      1141 NLLSIVMLTVDVTFGMALWDISYNAVSLINLVSAGVMSVEFVSHITRSFAISTKETWLER 1200

Qy      1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLIITLLGLHGLVFLPV 1260
Db      1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIFFFRNLNLIITLLGLHGLVFLPV 1260

Qy      1261 ILSYVGPDPNPALALEQKRAEEAVAAVMVASCPNHPRSVSTADNIYVNHSPGSIKGAGA 1320
Db      1261 ILSYVGPDPNPALALEQKRAEEAVAAVMVASCPNHPRSVSTADNIYVNHSPGSIKGAGA 1320

Qy      1321 ISNPLPNNGRQF 1332
Db      1321 ISNPLPNNGRQF 1332

RESULT 4
US-10-621-758A-4
; Sequence 4, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NEC1LI (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-4

Query Match      100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db      1 MAEAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60

Qy      61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db      61 TPARKITGDHLILLOKICPRLYTGNPTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120

Qy      121 VNLCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180
Db      121 VNLCHNTCSPNQSLFINVTRVAQIAGAGQLPAVVAEYAFYQHSFAEQSDVSCSRVRPAA 180
```

Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
Qy 181 ATLAGVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Db 181 ATLAGVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Qy 241 ARCNSQGGDDVATCSQDCCAASCPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDDVATCSQDCCAASCPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSY 420
Qy 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQVMSPEAQNRISLQDICYAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQVMSPEAQNRISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCWADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLMEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLMEAFLEEM 600
Qy 601 RAPQRMAGHFWQVTFATERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSYWSRV 660
Db 601 RAPQRMAGHFWQVTFATERSLEDEINRTTAEDLPFATSYIVIFLYISALGYSYWSRV 660
Qy 661 MVDKATLGLGGVAVVLGAVMAAGPFSYLGIRSSIVILQVVPFLVLSVGAQNRIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGPFSYLGIRSSIVILQVVPFLVLSVGAQNRIFVLE 720
Qy 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSSEALCFFLGALTMPAVRTFALTISGLAV 780
Db 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSSEALCFFLGALTMPAVRTFALTISGLAV 780
Qy 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCVKQELPPPGQEGGLLGFQKAYAPFL 840
Qy 841 LHWITRGVULLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFEVGA 900
Db 841 LHWITRGVULLFLALFGVLSYMSCHISVGLDQELALPKDSVLLDYFLFLNRYFEVGA 900
Qy 901 YFVTTILGYNFSBAGMNAICSSAGCNFSTQKIQVATEFPQSYLAIAPASSWVDDFDIW 960
Db 901 YFVTTILGYNFSBAGMNAICSSAGCNFSTQKIQVATEFPQSYLAIAPASSWVDDFDIW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRPSVSEQFKHLPWFNDPR 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKMSITMGSVRPSVSEQFKHLPWFNDPR 1020
Qy 1021 NIKCPKGLAAVSTSNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANTADL 1080
Db 1021 NIKCPKGLAAVSTSNLTSQGVLASRPMAYHKPLKNSQDYTEALRAARELAANTADL 1080
Qy 1081 RKVPGTDPAFEVFPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGLL 1140
Db 1081 RKVPGTDPAFEVFPYITVNFVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGGLL 1140
Qy 1141 NLLSIVMILVDTVGFPMALWDISTNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFPMALWDISTNAVSLINLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
Qy 1201 AKEATISMGSAVAGVAMTNLPGILVLGAKAQLOIIPFRNLNLTITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVAGVAMTNLPGILVLGAKAQLOIIPFRNLNLTITLGLLHGLVFLPV 1260

Qy 1261 ILSYVGPDPVPALALEQKRAEBAVAAMVASCPNHPFSRSTADNIYVNHSPFGSIKGAGA 1320
Db 1261 ILSYVGPDPVPALALEQKRAEBAVAAMVASCPNHPFSRSTADNIYVNHSPFGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332

RESULT 5

US-10-646-301A-4
; Sequence 4, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKPELSGSLMTLSNVCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHOPGYCAFYDECGKPELSGSLMTLSNVCLSN 60
Qy 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAQSQSDSCSRVRPAA 180
Qy 181 ATLAGVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Db 181 ATLAGVTMCVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQLNEGV 240
Qy 241 ARCNSQGGDDVATCSQDCCAASCPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDDVATCSQDCCAASCPATARPQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTLTTDPVELMSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSSY 420
Qy 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQVMSPEAQNRISLQDICYAPLNPDNT 480
Db 421 RYDSSLGPKNFGSGLDLDLLELLELQERLHLQVMSPEAQNRISLQDICYAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTPKDGFTALAL 540

QY 541 SCWADYGAVPFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
DB 541 SCWADYGAVPFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
QY 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
DB 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTFALTSGLA 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGGFFQKAYAPFL 840
QY 841 LHMWITRGVVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMWITRGVVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSFTOKI QYATEFPQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSFTOKI QYATEFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFINDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPWFINDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPNKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPNKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAREVPPYITNVFVEQYLTILPEGLFMLSCLGVPTFAVSCLLGLDLSGGL 1140
DB 1081 RKVPGTDPAREVPPYITNVFVEQYLTILPEGLFMLSCLGVPTFAVSCLLGLDLSGGL 1140
QY 1141 NLSIWMILVDTVGFMAWDISVNAVSLNLSAVGMSVEFVSHIRSPAISTKPTWLER 1200
DB 1141 NLSIWMILVDTVGFMAWDISVNAVSLNLSAVGMSVEFVSHIRSPAISTKPTWLER 1200
QY 1201 AKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFPFRMLLTITLGLHLGVFLPV 1260
DB 1201 AKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFPFRMLLTITLGLHLGVFLPV 1260
QY 1261 ILSYVGPVNPALAEQKRAEEAAVAVMVASCPNHPRSYSTADNIYVNHSPFGSIRKAGA 1320
DB 1261 ILSYVGPVNPALAEQKRAEEAAVAVMVASCPNHPRSYSTADNIYVNHSPFGSIRKAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 6
US-10-663-208A-4
; Sequence 4, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J061603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; PRIOR FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-663-208A-4

Query Match 100.0%; Score 6909; DB 36; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAEAGRLGWLWALLRLLAQSEPTTTHQPGYCAFYDECGRNPELSSGLMTLSNVCSLSN 60
DB 1 MAEAGRLGWLWALLRLLAQSEPTTTHQPGYCAFYDECGRNPELSSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSLFINVTRVAOLGAGQLPAVVAEYAFYOHSPABQSYDSCSRVRVPA 180
DB 121 VNLHCHNTCSNQSLFINVTRVAOLGAGQLPAVVAEYAFYOHSPABQSYDSCSRVRVPA 180
QY 181 ATLAGVTGCVGYGSALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATLAGVTGCVGYGSALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCHESQDDVATCSCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCHESQDDVATCSCQCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLQOFFQGWGTWASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLQOFFQGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHGFPPFTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTDPVELWSAPNSQARSEKAFHQHGFPPFTNQVILTAPNRSY 420
QY 421 RYDLSLLGPKNPSGILDLDLLELELQERLHLQVWSPQARNSLODICVAPLNPDNT 480
DB 421 RYDLSLLGPKNPSGILDLDLLELELQERLHLQVWSPQARNSLODICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKDGTAAL 540
QY 541 SCWADYGAVPFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
DB 541 SCWADYGAVPFPFLAIGYKGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLSEM 600
QY 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
DB 601 RAFORMMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISALGSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE 720
DB 661 MYDSKATLGLGGVAVVILGAVMAAMGFFSYLGRSSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTFALTSGLA 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGLALTMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPPGQGGGLLGGFFQKAYAPFL 840
QY 841 LHMWITRGVVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFEVGAPV 900


```
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAVSTSVNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAVSTSVNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYTTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVFPYTTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Qy 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLOIIFFRNLNLTITLGLLHGLVFLPV 1260
Db 1201 AKEATISGMSAVFAGVAMTNLPGLVLGLAKAQLOIIFFRNLNLTITLGLLHGLVFLPV 1260
Qy 1261 ILSYVGPDPNPALALEQKAEAAVAVMVASCPNHPSTADNIYNHSPFSGSIKGAGA 1320
Db 1261 ILSYVGPDPNPALALEQKAEAAVAVMVASCPNHPSTADNIYNHSPFSGSIKGAGA 1320
Qy 1321 ISNPLPNNRQF 1332
Db 1321 ISNPLPNNRQF 1332
```

RESULT 7

```
US-10-736-769-4
; Sequence 4, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-736-769-4

```
Query Match 100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
```

```
Db 61 TPARKITGDHLILLQKICPRLYTGNTQACCSAKQLVLSLEASLITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVVPAA 180
Db 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFARQSYDSCSRVVPAA 180
Qy 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQQAQVSGIQPLNEGV 240
Db 181 ATLVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPQQAQVSGIQPLNEGV 240
Qy 241 ARCNSQGGDDVATCSCODCAASCPALARQALDSTFYLGQMPGSLVLIILCSVPAVVTI 300
Db 241 ARCNSQGGDDVATCSCODCAASCPALARQALDSTFYLGQMPGSLVLIILCSVPAVVTI 300
Qy 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTILGQFFQCGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSNVDPKGTSLSDKLSFSTHTILGQFFQCGWGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPGPPFRTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPGPPFRTNQVILTAPNRSSY 420
Qy 421 RYDSILLGPKNPSGILDLDDLELLELOERLRLHQLVMSPEAQRNITSLODICYAPLNPDNT 480
Db 421 RYDSILLGPKNPSGILDLDDLELLELOERLRLHQLVMSPEAQRNITSLODICYAPLNPDNT 480
Qy 481 SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDKWDHFLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDKWDHFLYCANAPLTPKDGFTALAL 540
Qy 541 SCWADYGAPVPPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCWADYGAPVPPFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM 600
Qy 601 RAPQRMAGMPQVTTAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGSSYSWSRV 660
Db 601 RAPQRMAGMPQVTTAERSLEDEINRTTAEDLPITPATSYIIVIFLYISIALGSSYSWSRV 660
Qy 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGDADNIFVLVE 720
Db 661 MYDSKATILGLGVAVVLGAVMAAGFFSYLGRSSILVILQVVPFLVLSVGDADNIFVLVE 720
Qy 721 YORLPRRPGEPREVHIGRALGRVAPSMLLCSLEAICPFGLGALTMPAVRTTALTSLGLAV 780
Db 721 YORLPRRPGEPREVHIGRALGRVAPSMLLCSLEAICPFGLGALTMPAVRTTALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPPCQGGEGILLGPFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQBELPPPCQGGEGILLGPFQKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSEAGNNAICSSAGCNFSTFKIQIYATEFPQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVNSLCLKNCMSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAVSTSVNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAVSTSVNLTSDGOVLASRFMAVHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAFEVFPYTTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Db 1081 RKVPGTDPAFEVFPYTTITNVFEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
Qy 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
Db 1141 NLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTWLER 1200
```

```
QY 1201 AKEATISMSGSAVFAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLLITLLGLLHGLVFLPV 1260
DB 1201 AKEATISMSGSAVFAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLLITLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYNVHSPSGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYNVHSPSGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 8
US-10-750-386-4
; Sequence 4, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Galvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujjainwalla, Feroze
; APPLICANT: Altmann, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONISTS
; FILE REFERENCE: A36104 074669.0103
; CURRENT APPLICATION NUMBER: US/10/750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; TYPE: PRT
; LENGTH: 1332
; ORGANISM: Homo sapiens
US-10-750-386-4

Query Match 100.0%; Score 6909; DB 37; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLGRWLLALLRLAQSEPYTTHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLGRWLLALLRLAQSEPYTTHQPGYCAFYDECKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGQLPAVVAEAYQHSFABQSDYDSCSRVRVPA 180
DB 121 VNLHCHNTCSPNOSLFINVTRVAQLGAGQLPAVVAEAYQHSFABQSDYDSCSRVRVPA 180
QY 181 ATLAVTGTCVGSALCNQRLWLFQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
DB 181 ATLAVTGTCVGSALCNQRLWLFQDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
QY 241 ARCNEGQDDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVFVAVVTI 300
DB 241 ARCNEGQDDVATCSQDCAACPAIARPOALDSTFYLQMPGSLVLIILCSVFVAVVTI 300
QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFPQGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFPQGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLETTDPVELSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
DB 361 IPVVALAAGLVFTLETTDPVELSAPNSQARSEKAFHDQHFPGFFRTNOVILTAPNRSY 420
QY 421 RYDSLILGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPNDT 480
DB 421 RYDSLILGPKNFSGILDLDLLELLELQERLRLHQLVMSPEAQRNISLQDICYAPLNPNDT 480
```

```
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTLAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTLAL 540
QY 541 SCWADYCAPVPPFLAIGGYKGKDYSEABALIMTFTSLNNYPAGDPRLAQAKLWBEAFLEEM 600
DB 541 SCWADYCAPVPPFLAIGGYKGKDYSEABALIMTFTSLNNYPAGDPRLAQAKLWBEAFLEEM 600
QY 601 RAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSSMSRV 660
DB 601 RAFQRRMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSSYSSMSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVSGADNIFIFVLE 720
DB 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGIRSLVILQVVPFLVLSVSGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRFTALTSLAV 780
DB 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRFTALTSLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVKPQELPPPGQEGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVKPQELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHMTRGVVLLFLALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHMTRGVVLLFLALFGVSLYSMCHISVGLDQBLALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCSITMGSVRPSVEQFHKYLFWFLNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDQVLSAFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGGLAAYSTSVNLTSDQVLSAFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAPFVFPYTIITNVFYEQLYTLPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
DB 1081 RKVPGTDPAPFVFPYTIITNVFYEQLYTLPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
QY 1141 NLLSVMLVDTVGFMAWDISYNAVSLINLSVAGMSVEPVSHITRGSFAISTKPTWLER 1200
DB 1141 NLLSVMLVDTVGFMAWDISYNAVSLINLSVAGMSVEPVSHITRGSFAISTKPTWLER 1200
QY 1201 AKEATISMSGSAVFAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLLITLLGLLHGLVFLPV 1260
DB 1201 AKEATISMSGSAVFAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLLITLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYNVHSPSGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYNVHSPSGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 9
PCT-US01-04098A-1831
; Sequence 1831, Application PC/TUS0104098A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: PCT/US01/04098A
; CURRENT FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-04098A-1831

Query Match          99.8%; Score 6896; DB 1; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MABAGLRGWLWALLRLAQSPYTTIHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MABAGLRGWLWALLRLAQSPYTTIHOPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
DB 121 VNLHCHTCSNQSLFINTRVAQLGAGQLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLVGTMCGVGSALCNQRLNFGDGTGNGLAPLIDITFHLLERQAVGSGIQLNEGV 240
DB 181 ATLVGTMCGVGSALCNQRLNFGDGTGNGLAPLIDITFHLLERQAVGSGIQLNEGV 240
QY 241 ARCNESSQGDVATCSQDCAASCALARPQALDSTFYLQMPGSLVLIILCSFVAVVTI 300
DB 241 ARCNESSQGDVATCSQDCAASCALARPQALDSTFYLQMPGSLVLIILCSFVAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHFGPPFRTNQVILTAPESSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDOHFGPPFRTNQVILTAPESSY 420
QY 421 RYDLSLLGPKNFGSLIDLDLLELLELOERLHLQVWSPQARNISLODICYAPLNPDNT 480
DB 421 RYDLSLLGPKNFGSLIDLDLLELLELOERLHLQVWSPQARNISLODICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTMGQTSQVDWKDHFLYCANAPLTKFGQTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTMGQTSQVDWKDHFLYCANAPLTKFGQTALAL 540
QY 541 SCHADYGAPVFPFLAIGGKGYDSEAEALIMTFSLNYPAGDPRLAQKLEWAEFLBEM 600
DB 541 SCHADYGAPVFPFLAIGGKGYDSEAEALIMTFSLNYPAGDPRLAQKLEWAEFLBEM 600
QY 601 RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSWSRV 660
DB 601 RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGYSWSRV 660
QY 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLIGIRSSLVLQVVPFLVLSVGADNIFVLE 720
```

```
DB 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLIGIRSSLVLQVVPFLVLSVGADNIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEACPFPLGALTMPAVRTTALTSGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEACPFPLGALTMPAVRTTALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKRQESRLDYCCCKVQBELPPPGQEGGLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKRQESRLDYCCCKVQBELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHMITRGVVLILFLALFGVSLYSMCHI SVGLDOELALPKDSYLLDPLFLNRYFVGAPV 900
DB 841 LHMITRGVVLILFLALFGVSLYSMCHI SVGLDOELALPKDSYLLDPLFLNRYFVGAPV 900
QY 901 YFVVTILGYNFSSEAGNNAICSSAGCNPFSTQKIYVATEFPQOSYLAI PASSWVDDPIDW 960
DB 901 YFVVTILGYNFSSEAGNNAICSSAGCNPFSTQKIYVATEFPQOSYLAI PASSWVDDPIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRSPVSEQFHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSITMGSVRSPVSEQFHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQVTLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTVNLTSQVTLASRFMAHYHKLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFVPPYTTITNVFYEQYLTILPEGLFMLCLVPTFAVSCILLGLDLRSGLL 1140
DB 1081 RKVPGTDPAFVPPYTTITNVFYEQYLTILPEGLFMLCLVPTFAVSCILLGLDLRSGLL 1140
QY 1141 NLLSVIMILVDVTGFMALWDISYNAVSLINLVSAYGMSVEFYSHITRSFAISTKPTWLER 1200
DB 1141 NLLSVIMILVDVTGFMALWDISYNAVSLINLVSAYGMSVEFYSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISGMSAVFAGVAMTNLPGIILVLGLAKAQLIQIPFRLNLITLGLLHGLVFLPV 1260
DB 1201 AKEATISGMSAVFAGVAMTNLPGIILVLGLAKAQLIQIPFRLNLITLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIVVNSPBGSIKGAGA 1320
DB 1261 ILSYVGPDPVNPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIVVNSPBGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332
```

```
RESULT 10
US-10-239-316-9
; Sequence 9, Application US/10239316
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiyama
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703USOP
; CURRENT APPLICATION NUMBER: US/10/239,316
; CURRENT FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-10-239-316-9
```

```
Query Match          99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Qy 61 TPARKITGDHILILLOKICPRLYTGNTQACCSAQVLSLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILILLOKICPRLYTGNTQACCSAQVLSLEASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVVAEAFYQHSFAEQSYDSCSRVRVPAA 180
Qy 181 ATLVGTCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLBPQAVGSGIQPLNEGV 240
Db 181 ATLVGTCGVYGSALCNAQRWLNFGQDTGNGLAPLDITFHLLBPQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSCQDCACPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSCQDCACPAIARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LVLGFRVAPARDKSKVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLVS 360
Db 301 LVLGFRVAPARDKSKVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLVS 360
Qy 361 IPVVALAAGLVFTELTTPDELVELWSPNQARSEKAFHQHGFPPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELVELWSPNQARSEKAFHQHGFPPFRTNQVILITAPNRSY 420
Qy 421 RYDLSLLGPNFSGILDLDLLELELOERLHLQVMSPEAQRNISLODICVAPLNPDNT 480
Db 421 RYDLSLLGPNFSGILDLDLLELELOERLHLQVMSPEAQRNISLODICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTPKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTPKQGTALAL 540
Qy 541 SCHADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEAFLEEM 600
Db 541 SCHADYGAVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEAFLEEM 600
Qy 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPFATSYIYIFLYISALGSYSWSRV 660
Db 601 RAFORMMAGMFQVTFPABRSLEDEINRTTAEDLPFATSYIYIFLYISALGSYSWSRV 660
Qy 661 MVDSKATLGLGGVAVVLGAVMAAMGPFYSYLGIRSSLVILQVVPFLVLSVGAADNIFVLE 720
Db 661 MVDSKATLGLGGVAVVLGAVMAAMGPFYSYLGIRSSLVILQVVPFLVLSVGAADNIFVLE 720
Qy 721 YQRLPRPGPREVHIGRALGRVAPSMILCSISEATCFELGALTPMPAVRTALTSLGLAV 780
Db 721 YQRLPRPGPREVHIGRALGRVAPSMILCSISEATCFELGALTPMPAVRTALTSLGLAV 780
Qy 781 ILDFLLQMAFAVALLSLDKRQASRLDCCCVPQELPPPGQEGGLLGGFFQKAYAPFL 840
Db 781 ILDFLLQMAFAVALLSLDKRQASRLDCCCVPQELPPPGQEGGLLGGFFQKAYAPFL 840
Qy 841 LHMWTRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVFGAPV 900
Db 841 LHMWTRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVFGAPV 900
Qy 901 YFVTTILGYNFSSAGWNAICSSAGCNFSEFTOKIQATEPPEOSYLAIPASSWVDDPIDW 960
Db 901 YFVTTILGYNFSSAGWNAICSSAGCNFSEFTOKIQATEPPEOSYLAIPASSWVDDPIDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKWMSITMGSVRPSVQFHKYLFWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSVTNSLNCNKWMSITMGSVRPSVQFHKYLFWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDGQVLSRFRMAYHKPLKNSQDYTEALRAARELANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSDGQVLSRFRMAYHKPLKNSQDYTEALRAARELANITADL 1080

Qy 1081 RKVEGTDPAFEVFPYTTITNVFYEQYLITLPEGLFMLSLCLVPTPAVSCLLILGLDLRSGLL 1140
Db 1081 RKVEGTDPAFEVFPYTTITNVFYEQYLITLPEGLFMLSLCLVPTPAVSCLLILGLDLRSGLL 1140
Qy 1141 NLLSIVMLVDTVGFPMALWDISYNAVSLINLVSAGVMSVEFVSHITSPASTKPTWLER 1200
Db 1141 NLLSIVMLVDTVGFPMALWDISYNAVSLINLVSAGVMSVEFVSHITSPASTKPTWLER 1200
Qy 1201 AKRATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFFFRLNLLITLLGLLHGLVFLPV 1260
Db 1201 AKRATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFFFRLNLLITLLGLLHGLVFLPV 1260
Qy 1261 ILSVGVDPNPALEQKRAEEAAVAAVMVASCNPHPSRVSTADNIYVNHSPSGSIKGAGA 1320
Db 1261 ILSVGVDPNPALEQKRAEEAAVAAVMVASCNPHPSRVSTADNIYVNHSPSGSIKGAGA 1320
Qy 1321 ISNFLPNNGRQF 1332
Db 1321 ISNFLPNNGRQF 1332
RESULT 11
US-10-258-899A-1831
; Sequence 1831, Application US/10258899A
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunqing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258,899A
; PRIOR FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/554,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT

ORGANISM: Homo sapiens
US-10-258-899A-1831

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSPEYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSPEYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLLILLOKI CRLTYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKI CRLTYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120

QY 121 VNLHCHNTSPNQSFLINTRVAQLGAGQLPAVYAEAFYQHSFAEQSDSCSRVRPAA 180
DB 121 VNLHCHNTSPNQSFLINTRVAQLGAGQLPAVYAEAFYQHSFAEQSDSCSRVRPAA 180

QY 181 ATLAVTGTCGVGSGALCNAQRWLNFGQDTGNGLAPLIDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAVTGTCGVGSGALCNAQRWLNFGQDTGNGLAPLIDITFHLLPEQAVGSGIQPLNEGV 240

QY 241 ARCNSQGDVATCSCODCAASCPTARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNSQGDVATCSCODCAASCPTARPOALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

QY 301 LVLGFRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLVS 360
DB 301 LVLGFRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLVS 360

QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHQHFGPFRTNQVILTAPNRSY 420

QY 421 RYDLSLLGPNKSGILDLLLELLELQRLHRLQWSPEAQRNLSQDICVAPLNPDNT 480
DB 421 RYDLSLLGPNKSGILDLLLELLELQRLHRLQWSPEAQRNLSQDICVAPLNPDNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVMDKDFLYCANAPLTPKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGQTSQVMDKDFLYCANAPLTPKDGTALAL 540

QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAALWEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALINTFSLNNYPAGDPRLAQAALWEAFLEEM 600

QY 601 RAPORMAGMFOVTTAERSLDEINRTAEDLPATSYIVIFLYISLALGSYSWSRV 660
DB 601 RAPORMAGMFOVTTAERSLDEINRTAEDLPATSYIVIFLYISLALGSYSWSRV 660

QY 661 MYDSKATLGLGVAVVLGVMAAMGFFSYLGRSSVLVLOVVPFLVLSVGDNIPIFVLE 720
DB 661 MYDSKATLGLGVAVVLGVMAAMGFFSYLGRSSVLVLOVVPFLVLSVGDNIPIFVLE 720

QY 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTTPMPAVRTEALTGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSLSEALCFPLGALTTPMPAVRTEALTGLAV 780

QY 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCKVQELPPPPQGBGLLIGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKRQESRLDVCCKVQELPPPPQGBGLLIGFFQKAYAPFL 840

QY 841 LHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELPKDSYLLDYFLFLNRYFVGVAPV 900
DB 841 LHWITRGVWVLLFLALFGVSLYSMCHISVGLDQELPKDSYLLDYFLFLNRYFVGVAPV 900

QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSTOKIQYATEFFPQSYLAIPASSWVDQFDW 960
DB 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSTOKIQYATEFFPQSYLAIPASSWVDQFDW 960

QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQFHKYLFWFLNDRP 1020

DB 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNKMSITMGSVRPSVEQFHKYLFWFLNDRP 1020

QY 1021 NIKPRKGGLAAYSTVNLTSQVTLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKPRKGGLAAYSTVNLTSQVTLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080

QY 1081 RKVPGTDPAPFVPPYTTINVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140
DB 1081 RKVPGTDPAPFVPPYTTINVFYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLL 1140

QY 1141 NLLSVIMILVDTVGFMAWDISYNAVSLINLVSACMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSVIMILVDTVGFMAWDISYNAVSLINLVSACMSVEFVSHITRSFAISTKPTWLER 1200

QY 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIIFFRNLNLLITLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIIFFRNLNLLITLGLLHGLVFLPV 1260

QY 1261 ILSYVGPDVNPALALEQKRAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFEGSIKAGA 1320
DB 1261 ILSYVGPDVNPALALEQKRAEBAVAAVMVASCPNHPRSVSTADNIYVNHSPFEGSIKAGA 1320

QY 1321 ISNPLPNNGRQF 1332
DB 1321 ISNPLPNNGRQF 1332

RESULT 12
US-10-293-244-1831
; Sequence 1831, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1831

Query Match 99.8%; Score 6896; DB 32; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAEAGLRGWLWALLRLAQSPEYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSPEYTHIQPGYCAFYDECGKPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLLILLOKI CRLTYTGPNTQACCSAKQLVSLASISITKALLTRCPACSDNF 120

Db 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSEASLSITTKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSIFINVTRVAQIAGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSIFINVTRVAQIAGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAGVTGCVGVGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEG 240
Db 181 ATLAGVTGCVGVGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEG 240
Qy 241 ARCNEQGDVATCSQDCAASCPAIAARQALDSTFYLGOMPGSLVLIILCSVFVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIAARQALDSTFYLGOMPGSLVLIILCSVFVVTI 300
Qy 301 LLVGRVAPADKSKWDPKGTSDKLSFSTHTLLQFQGWGTWASWPLTILVLVS 360
Db 301 LLVGRVAPADKSKWDPKGTSDKLSFSTHTLLQFQGWGTWASWPLTILVLVS 360
Qy 361 IPVVALAAGLVFTELTTPDELWSAPNSOARSEKAFHDOHFGPPRRTNOVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPDELWSAPNSOARSEKAFHDOHFGPPRRTNOVILTAPNRSY 420
Qy 421 RYDLSLLGPKNPSGLDLDLLELLELQERLRLHQLQWSPQARQNSLQDICYAPLNPNT 480
Db 421 RYDLSLLGPKNPSGLDLDLLELLELQERLRLHQLQWSPQARQNSLQDICYAPLNPNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPFLYCANAPLTPKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHPFLYCANAPLTPKQGTALAL 540
Qy 541 SCWADYGAPVFPFLTAIGYKGDYSEALIMTFSLNNYPAGDPRLAQAKWEFAFLEM 600
Db 541 SCWADYGAPVFPFLTAIGYKGDYSEALIMTFSLNNYPAGDPRLAQAKWEFAFLEM 600
Qy 601 RAFORMMAGMFOVTTAERSLDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660
Db 601 RAFORMMAGMFOVTTAERSLDEINRTTAEDLPFATSYIVIFLYISIALGYSWSRV 660
Qy 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVLQVVPFLVLSVGADNIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSILVLQVVPFLVLSVGADNIFVLE 720
Qy 721 YQRLPRRPEPREVHIGRALGRVAPSMILCSSEALCFPLGALTTPMPAVRTEALTSGLA 780
Db 721 YQRLPRRPEPREVHIGRALGRVAPSMILCSSEALCFPLGALTTPMPAVRTEALTSGLA 780
Qy 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQCGRGLLIGFPKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPPQCGRGLLIGFPKAYAPFL 840
Qy 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Db 841 LHWITRGVVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
Qy 901 YFVTTILGYNFSRAGNATCSAGCNCFSTOKIYATFEPQSYLAIPASSWDDDFIDW 960
Db 901 YFVTTILGYNFSRAGNATCSAGCNCFSTOKIYATFEPQSYLAIPASSWDDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKMCSITMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCIKMCSITMGSVRPSVEQFHXYLPWFINDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Db 1021 NIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPLKNSQDYTEALRAARELAANITADL 1080
Qy 1081 RKVPGTDPAPEFPFYITNVFEOYLTILPEGLFMLSCLVPTFAVSCLLGLDURSGLL 1140
Db 1081 RKVPGTDPAPEFPFYITNVFEOYLTILPEGLFMLSCLVPTFAVSCLLGLDURSGLL 1140
Qy 1141 NLLSIWMLVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200
Db 1141 NLLSIWMLVDTVGFNALWDISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLER 1200

Qy 1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIOPFFRNLNLLITLGLLHGLVFLPV 1260
Db 1201 AKEATISMGSAVFAGVAMTNLPGLVLGLAKAQLIQIOPFFRNLNLLITLGLLHGLVFLPV 1260
Qy 1261 ILSVVGPDVNPALALEQKRAEEAAVAVMVASCPNHPRSRVSTADNIYVNHSPGSIKGAGA 1320
Db 1261 ILSVVGPDVNPALALEQKRAEEAAVAVMVASCPNHPRSRVSTADNIYVNHSPGSIKGAGA 1320
Qy 1321 ISNFLPNNGROF 1332
Db 1321 ISNFLPNNGROF 1332

RESULT 13
US-11-218-141-1831
; Sequence 1831, Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1831
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1831

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Fred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MAEAGLRGWLWALLRLLAQSEPYTTHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSEPYTTHQPGYCAFVDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSEASLSITTKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLQKICPRLYTGNTQACCSAKQLVSEASLSITTKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSNQSIFINVTRVAQIAGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSIFINVTRVAQIAGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Qy 181 ATLAGVTGCVGVGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEG 240
Db 181 ATLAGVTGCVGVGSALCNAQRWLNFGQDTGNGLAFLDITFHLLRPGQAVGSGIQPLNEG 240
Qy 241 ARCNEQGDVATCSQDCAASCPAIAARQALDSTFYLGOMPGSLVLIILCSVFVVTI 300
Db 241 ARCNEQGDVATCSQDCAASCPAIAARQALDSTFYLGOMPGSLVLIILCSVFVVTI 300


```
QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
QY 421 RYDSSLGPKNFGSILDLDDLLLELELQERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
DB 421 RYDSSLGPKNFGSILDLDDLLLELELQERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFL 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFL 600
QY 601 RAFQRMAGMFQVTFWAEASLEDEINRTTAEDLPIFATSVIVFLYISALAGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFWAEASLEDEINRTTAEDLPIFATSVIVFLYISALAGSYSSWSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAAGPFSYIGIRSSLVILQVVPFLVSVGADNIPFVLE 720
DB 661 MVDKATLGLGGVAVVLGAVMAAGPFSYIGIRSSLVILQVVPFLVSVGADNIPFVLE 720
QY 721 YQRLPRRPEPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPEPREVHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQBSRLDVCCKVQELPPPGQEGLLGCFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQBSRLDVCCKVQELPPPGQEGLLGCFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSTQKIQVATEFPQGSYLAIPASSWVDDFDIW 960
DB 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSTQKIQVATEFPQGSYLAIPASSWVDDFDIW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNKCMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSVNLTSDGQVLSRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSVNLTSDGQVLSRPMAYHKPLKNSQDYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEPVPPYITVNFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
DB 1081 RKVPGTDPAFEPVPPYITVNFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLL 1140
QY 1141 NLLSIVMILVDVTGFMALWDISVNASLNLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
DB 1141 NLLSIVMILVDVTGFMALWDISVNASLNLNLSAVGMSVEFVSHITRSFAISTKPTWLER 1200
QY 1201 AKEATISMGSAFVAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLTLLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAFVAGVAMTNLPGLVLGLAKAQLIQIFPFRNLNLTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPVNPALALEQKAEBAVAAVMVASCPNHPRSYSTADNIYVNHSPFGSINGAGA 1320
DB 1261 ILSYVGPVNPALALEQKAEBAVAAVMVASCPNHPRSYSTADNIYVNHSPFGSINGAGA 1320
QY 1321 ISNLPNNNGROF 1332
DB 1321 ISNLPNNNGROF 1332
```

RESULT 14

```
US-11-242-459-9
; Sequence 9, Application US/11242459
; GENERAL INFORMATION:
; APPLICANT: TANIYAMA, Yoshio
; APPLICANT: KITA, Shunbun
; APPLICANT: SATOMI, Tomoko Komiya
; TITLE OF INVENTION: Novel Protein, Process for Producing The Same And Use Thereof
; FILE REFERENCE: 2703US0P
; CURRENT APPLICATION NUMBER: US/11/242,459
; CURRENT FILING DATE: 2005-10-03
; PRIOR APPLICATION NUMBER: US/10/239,316
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: PCT/JP01/02279
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: JP2000-088595
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 59
; SEQ ID NO 9
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Human
US-11-242-459-9
```

Query Match 99.8%; Score 6896; DB 42; Length 1332;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCAPFYDECGKNPELSGLMTLSNVSLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIHPGVCAPFYDECGKNPELSGLMTLSNVSLSN 60
QY 61 TPARKITGDLHLLQKICPRLYTGPNTOACCQAKQLVSEASISITKALLTRCPACSDNP 120
DB 61 TPARKITGDLHLLQKICPRLYTGPNTOACCQAKQLVSEASISITKALLTRCPACSDNP 120
QY 121 VNLHCHTSPNQSLFINVTRVAQAGAGQLPVAYEAFYQHSFAEQSDSCSRVPA 180
DB 121 VNLHCHTSPNQSLFINVTRVAQAGAGQLPVAYEAFYQHSFAEQSDSCSRVPA 180
QY 181 ATLVGTMCGVYSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGSQIPLNEGV 240
DB 181 ATLVGTMCGVYSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVSGSQIPLNEGV 240
QY 241 ARCNESQGDVATCSCODCAASCPAARQALDSIFYLQMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNESQGDVATCSCODCAASCPAARQALDSIFYLQMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDQHFQFPFRTNQVILTAPNRSY 420
QY 421 RYDSSLGPKNFGSILDLDDLLLELELQERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
DB 421 RYDSSLGPKNFGSILDLDDLLLELELQERLHLQVMSPEAQRNLSLQDICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFL 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFL 600
QY 601 RAFQRMAGMFQVTFWAEASLEDEINRTTAEDLPIFATSVIVFLYISALAGSYSSWSRV 660
DB 601 RAFQRMAGMFQVTFWAEASLEDEINRTTAEDLPIFATSVIVFLYISALAGSYSSWSRV 660
```



```
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTTILGYNFSSBAGNNAICSSAGCNFFSTQKIQVATEFPBQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSBAGNNAICSSAGCNFFSTQKIQVATEFPBQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFPWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLTVAILSPRLYSCTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCILLGLDLRSGLNLLSIYVILVDTVGFMAWDISYNAVSLNLYVS 1173
Db 1141 FMLSCLVPTFAVSCILLGLDLRSGLNLLSIYVILVDTVGFMAWDISYNAVSLNLYVS 1200
Qy 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFPFRNLNLLITLGLLHGLVFLPVILSYVGPVDPNPALEQKABEAAVAVMSACP 1293
Db 1261 LIQIFPFRNLNLLITLGLLHGLVFLPVILSYVGPVDPNPALEQKABEAAVAVMSACP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNFPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNFPNNGRQF 1359

RESULT 16
PCT-US03-22467-44
; Sequence 44, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-22467-44

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLAQSBPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLAQSBPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHTCSNQSLFINTRVAQLGAGOLPAVAYEAFYQHSFABQSDVSCSRVRVPA 180
Db 121 VNLHCHTCSNQSLFINTRVAQLGAGOLPAVAYEAFYQHSFABQSDVSCSRVRVPA 180
Qy 181 ATLAGVTGTCVGTGSALCAQRWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
Db 181 ATLAGVTGTCVGTGSALCAQRWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQPLNEGV 240
Qy 241 ARCNEQGGDDVATSCQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
```

```
Db 241 ARCNEQGGDDVATSCQDCAASCPAIPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Qy 301 LILVGPVAPARDKSMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LILVGPVAPARDKSMVDPKGTSLSDKLSFSFTHLLGQFFQCGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFGRFTNQVILTAPNRSSY 420
Db 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHFGRFTNQVILTAPNRSSY 420
Qy 421 RYDLSLLGPNKPSGILLDILLLELLELRLHQLWSPQAQRNLSLODICYAPLNDPNT 480
Db 421 RYDLSLLGPNKPSGILLDILLLELLELRLHQLWSPQAQRNLSLODICYAPLNDPNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHPLYCANAPLTPKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKHPLYCANAPLTPKDGFTALAL 540
Qy 541 SCHADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLEEM 600
Db 541 SCHADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLMEAEFLEEM 600
Qy 601 RAFORMAGMPOVTTFAERSLDEINRTTAEDLPIPATSYIYVIFLYISLALGSSYSWSRV 660
Db 601 RAFORMAGMPOVTTFAERSLDEINRTTAEDLPIPATSYIYVIFLYISLALGSSYSWSRV 660
Qy 661 WYDSKATILGLGVAVVILGAVMAAGFPFSLYGRSSILVILQVVPFLVLSYVADNIFIPVLE 720
Db 661 WYDSKATILGLGVAVVILGAVMAAGFPFSLYGRSSILVILQVVPFLVLSYVADNIFIPVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMLLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKROASRLDVCCCKVQOELPPPQGGEGLLGFPKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROASRLDVCCCKVQOELPPPQGGEGLLGFPKAYAPFL 840
Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLFLNRYFVGAPV 900
Qy 901 YFVTTILGYNFSSBAGNNAICSSAGCNFFSTQKIQVATEFPBQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSBAGNNAICSSAGCNFFSTQKIQVATEFPBQSYLAIPASSWVDDFIDW 960
Qy 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCCLKNCSITMGSVRPSVEQFHKYLFPWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLTVAILSPRLYSCTISAHCNLYLLDSASRFMAYH 1080
Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCILLGLDLRSGLNLLSIYVILVDTVGFMAWDISYNAVSLNLYVS 1173
Db 1141 FMLSCLVPTFAVSCILLGLDLRSGLNLLSIYVILVDTVGFMAWDISYNAVSLNLYVS 1200
Qy 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFPFRNLNLLITLGLLHGLVFLPVILSYVGPVDPNPALEQKABEAAVAVMSACP 1293
Db 1261 LIQIFPFRNLNLLITLGLLHGLVFLPVILSYVGPVDPNPALEQKABEAAVAVMSACP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFSGIKGAGAIISNFPNNGRQF 1332
```

Db 1321 NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF 1359

RESULT 17
PCT-US03-40113-44
; Sequence 44, Application PC/TUS0340113
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K3-WI
; CURRENT APPLICATION NUMBER: PCT/US03/40113
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40113-44

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFDCEGKQPELSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSPYTTIHQPGYCAFDCEGKQPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGNTQACCSAKOLVLEASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSFLINTRVAQAGQLPAWYAEAFYQHSFAEQSDSCSRVRPAA 180
Db 121 VNLHCHTCSNQSFLINTRVAQAGQLPAWYAEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATAVGTGCVGYSALCNQRMNFQDGTGNGLAPLDITFHLLEPGQAVGSIQPLNEGV 240
Db 181 ATAVGTGCVGYSALCNQRMNFQDGTGNGLAPLDITFHLLEPGQAVGSIQPLNEGV 240
QY 241 ARCNEQGDVATCSQDCCAAFCAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCCAAFCAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
QY 361 IPVALLAAGLVFTELTDVPELWSAENQARSEKAFHDQFGFPFRTNOVILTAPNRSY 420
Db 361 IPVALLAAGLVFTELTDVPELWSAENQARSEKAFHDQFGFPFRTNOVILTAPNRSY 420
QY 421 RYDSSLGPNFSGIILDLLELLEQLERHLQVWSPQAORNI SLQDICVAPLNPDNT 480
Db 421 RYDSSLGPNFSGIILDLLELLEQLERHLQVWSPQAORNI SLQDICVAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGFTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEBALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEBALIMTFSLNNYPAGDPRLAQAKLWEAFLEEM 600
QY 601 RAFORMMAGMFQVTFMAERSLEINRTTAEDLPFATSYIVIFLYISLALGSSWSRV 660

Db 601 RAFORMMAGMFQVTFMAERSLEINRTTAEDLPFATSYIVIFLYISLALGSSWSRV 660
QY 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MVDKATLGLGGVAVVLGAVMAAGFFSYLGRSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMCLLSLSAICFFLGCALTPMPAVRTPALTSGLAV 780
Db 721 YORLPRRPGEPREVIHGRALGRVAPSMCLLSLSAICFFLGCALTPMPAVRTPALTSGLAV 780
QY 781 ILDELLQMSAFVALLSLDSKQEAASRLDVCVCKPQELPPPGQEGGLLGFQKAYAPFL 840
Db 781 ILDELLQMSAFVALLSLDSKQEAASRLDVCVCKPQELPPPGQEGGLLGFQKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVLSYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
Db 841 LHWITRGVVLFLALFGVLSYNSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGPV 900
QY 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIYATEFPEQSYLAIPASSWVDDFIDW 960
Db 901 YFVTTLGVNFSSEAGMNAICSSAGCNPFSTQKIYATEFPEQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNCSITMGSVRPSVEQFHKYLWFLNDRP 1020
QY 1021 NIKCPKGLAAYSTVNLTSQGVLT-----ASRFMAVH 1053
Db 1021 NIKCPKGLAAYSTVNLTSQGVLTVAIILSPRLEYSGTISAHCNLYLLSASRFMAVH 1080
QY 1054 KPLKNSQDYTBALRAARELANITADLRKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1113
Db 1054 KPLKNSQDYTBALRAARELANITADLRKVPCTDPAFEVFPYTTINVFYEQYLTILPEGL 1113
QY 1114 FMLSCLVPTFAVSCLLGLDLRLSGLNLSIVMLVDTVGFMAWLDISYNAVSLINLVS 1173
Db 1114 FMLSCLVPTFAVSCLLGLDLRLSGLNLSIVMLVDTVGFMAWLDISYNAVSLINLVS 1173
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAGVAMTNLPGILVLGLAKAQ 1233
QY 1234 LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPALEQKRAEAAVAVVASC 1293
Db 1234 LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPALEQKRAEAAVAVVASC 1293
QY 1261 LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPALEQKRAEAAVAVVASC 1320
Db 1261 LIQIFPFRNLNLTLLGLLHGLVFLPVILSYVGPDPNPALEQKRAEAAVAVVASC 1320
QY 1294 NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPGSIKGAGAI SNFLPNNGRQF 1359

RESULT 18
PCT-US05-01469-44
; Sequence 44, Application PC/TUS0501469
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF
; FILE REFERENCE: 36134-PCT 074669.0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; CURRENT FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US 60/537,341
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-01469-44

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLASISYTKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLASISYTKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSNQSFLFINTVRAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSFLFINTVRAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180

QY 181 ATAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240

QY 241 ARCNESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

QY 301 LLVGFVRVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420

QY 421 RYDLSLLGPNFSGIILDLLELLELQBLRLHLQWSPEAQRNLSQDICVAPLNPDT 480
Db 421 RYDLSLLGPNFSGIILDLLELLELQBLRLHLQWSPEAQRNLSQDICVAPLNPDT 480

QY 481 SLYDCCINSILQYFQNNRTLLLTANQTLWGQTSQVDKQHLKYCANAPLTKQGTALAI 540
Db 481 SLYDCCINSILQYFQNNRTLLLTANQTLWGQTSQVDKQHLKYCANAPLTKQGTALAI 540

QY 541 SCHADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAPLEEM 600
Db 541 SCHADYGAPVFPFLAIGGYKGDYSEALIMTFSLNNYPAGDPRLAQAKLWEEAPLEEM 600

QY 601 RAFORMAGMPQVTFATRSLSDEINRTTAEDLPATSYIVIFLYISIALGYSWSRV 660
Db 601 RAFORMAGMPQVTFATRSLSDEINRTTAEDLPATSYIVIFLYISIALGYSWSRV 660

QY 661 WYDSKATLGLGVAVVILGAVMAAGPFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE 720
Db 661 WYDSKATLGLGVAVVILGAVMAAGPFSYLGIRSSLVILQVVPFLVSVGADNIFIFVLE 720

QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEALICPFLGALTMPAVRTEALTSGLA 780
Db 721 YQRLPRRPGEPREVHIGRALGRVAPSMILCSLSEALICPFLGALTMPAVRTEALTSGLA 780

QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKQQLPPCGEGLLGFPQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCVKQQLPPCGEGLLGFPQKAYAPFL 840

QY 841 LHWITRGVLLALFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFNRYFEVAPV 900
Db 841 LHWITRGVLLALFLALFGVLSYMSCHISVGLQDELALPKDSYLLDYFLFNRYFEVAPV 900

QY 901 YFVVTILGYNFSBAGNNAICSSAGCNFSTQKIYATFEPFQSYLAIPASSWDDFDIW 960
Db 901 YFVVTILGYNFSBAGNNAICSSAGCNFSTQKIYATFEPFQSYLAIPASSWDDFDIW 960

QY 961 LTFPSSCCRLYISGPNKDFCPSVTNSLNCNKMSTMGSVRPSVEQFHXYLPWFINDRP 1020
Db 961 LTFPSSCCRLYISGPNKDFCPSVTNSLNCNKMSTMGSVRPSVEQFHXYLPWFINDRP 1020

QY 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053
Db 1021 NIKCPKGLAAYSTSVNLTSDQVLT-----ASRFMAYH 1053

QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFVFPFYTTITNVFYEQLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFVFPFYTTITNVFYEQLTILPEGL 1140

QY 1114 FMLSCLIVPTFAVSCLLGLDLRLSGLNLNLSIVMILVDTVGFMALWDSINAVSLNLVS 1173
Db 1141 FMLSCLIVPTFAVSCLLGLDLRLSGLNLNLSIVMILVDTVGFMALWDSINAVSLNLVS 1200

QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260

QY 1234 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYGPVNPALALEQKEAEBAVAAMVVASCP 1293
Db 1261 LIQIPFFRLNLITLGLLHGLVFLPVILSYVYGPVNPALALEQKEAEBAVAAMVVASCP 1320

QY 1294 NHPSRVSTADNTVYVNHSPFGSIKGAGAINFNPNNGRQF 1332
Db 1321 NHPSRVSTADNTVYVNHSPFGSIKGAGAINFNPNNGRQF 1359

RESULT 19

PCT-US05-27579-22
; Sequence 22, Application PC/TUS0527579
; GENERAL INFORMATION:
; APPLICANT: Mount Sinai School of Medicine
; APPLICANT: Ioannou, Yiannis
; APPLICANT: Davies, Joanna P.
; TITLE OF INVENTION: NPC1L1 AND NPC1L1 INHIBITORS AND METHODS OF USE THEREOF
; FILE REFERENCE: 2201581-WO
; CURRENT APPLICATION NUMBER: PCT/US05/27579
; CURRENT FILING DATE: 2005-08-12
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-27579-22

Query Match 99.5%; Score 6872.5; DB 1; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTHQGYCAFYDECGKNPELSGLMTLSNVCSLSN 60

QY 61 TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLASISYTKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLQKICPRLYTGPNTQACCSAKQLVSLASISYTKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSNQSFLFINTVRAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180
Db 121 VNLHCHNTCSNQSFLFINTVRAQLGAGOLPAVVAEAFYQHSFASQSDSCSRVRPAA 180

QY 181 ATAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATAVGTMCGVYGSALCNAQRWLNFGDGTNGGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240

QY 241 ARCNESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNESQDDVATCSCODCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300

QY 301 LLVGFVRVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360
Db 301 LLVGFVRVAPARDKSKMVDPKGTSLSKLSFSFTHLLGQFFQGWGTWVASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTTPVELMSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSY 420

```
Db 361 IPVVALAAGLVFTELTTPDVELMSAPNSQARSEKAFHQHFGPPRPTNQVILTAPNRSY 420
QY 421 RYDSLLGPKNPGSGLDLDLLELLELQERLHLQWSPBAQRNLSLODICYAPLNPNT 480
Db 421 RYDSLLGPKNPGSGLDLDLLELLELQERLHLQWSPBAQRNLSLODICYAPLNPNT 480
QY 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGTSQVDMKDHFLYCANAPLTFKDTALAL 540
Db 481 SLYDCCINSLLQYFONNRTLLLTANQTLMGTSQVDMKDHFLYCANAPLTFKDTALAL 540
QY 541 SCWADYGAPVFFFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
Db 541 SCWADYGAPVFFFLAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAEFLBEM 600
QY 601 RAFORMMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISLAGSYSSWSRV 660
Db 601 RAFORMMAGMFOVTTAERSLEDEINRTTAEDLPITFATSYIVIFLYISLAGSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFESYLGIRSSLVILQVPPRLVLSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFESYLGIRSSLVILQVPPRLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVIHGRALGVAPSMMLCSLSEALCFPLGALTTPMPAVRTFALTSLGLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGVAPSMMLCSLSEALCFPLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKPQELPPPGQGGGLLGGFFQKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLSKQREASRLDVCCCKPQELPPPGQGGGLLGGFFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAAPV 900
Db 841 LHWITRGVVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAAPV 900
QY 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSFTQKIQVATEFPQSYLAIPASSWVDDFDW 960
Db 901 YFVTTLGYNFSSEAGNNAICSSAGCNFSFTQKIQVATEFPQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLCLKNCMSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
QY 1021 NIKCPKGGAAVSTSNLTSQGVLT-----ASRPMAYH 1053
Db 1021 NIKCPKGGAAVSTSNLTSQGVLT-----ASRPMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTTITNVFYEQLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEFVFPYTTITNVFYEQLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLLNLLSIVMILVDTVGMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLLNLLSIVMILVDTVGMALWDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPATSKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIFFRNLNLTLLGLLHGLVFLPVLVSYGPDVNPALAEQKAEAEAAVAVMSACP 1293
Db 1261 LIQIFFRNLNLTLLGLLHGLVFLPVLVSYGPDVNPALAEQKAEAEAAVAVMSACP 1320
QY 1294 NHPSRVSTADNIYVNHFSFGSIKAGAI SNFPLNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHFSFGSIKAGAI SNFPLNNGRQF 1359
```

RESULT 20

US-10-258-899A-1830

; Sequence 1830, Application US/10258899A

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

```
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Cao, Yicheng
; APPLICANT: Ma, Yunqing
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Xue, Aidong
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Goodrich, Ryle
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 787CIP2-2G/US
; CURRENT APPLICATION NUMBER: US/10/258,899A
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: PCT/US01/04098
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 09/774,434
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-258-899A-1830
```

Query Match 99.5%; Score 6872.5; DB 32; Length 1359;

Best Local Similarity 97.9%; Pred No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

```
QY 1 MAEAGLRGMLLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
Db 1 MAEAGLRGMLLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLOKICPRLYTGNTQAACSAKQLVSLASLSTIKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILLOKICPRLYTGNTQAACSAKQLVSLASLSTIKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAEQSDYDSRVRVPA 180
Db 121 VNLHCHNTCSNPQSLFINVTRVAQLGAGOLPAVVAEYAFYQHSFAEQSDYDSRVRVPA 180
QY 181 ATLA VGTMCVGYGSALCNAQRWLNFGDGTNGCLAPLDITFHLLPEQAVGSGIQPLNEGV 240
Db 181 ATLA VGTMCVGYGSALCNAQRWLNFGDGTNGCLAPLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDDVATCSCODCAACPAIRPAQALDSTFYLGOMPSGLVLIILCSVFAVVTI 300
Db 241 ARCNSQGGDDVATCSCODCAACPAIRPAQALDSTFYLGOMPSGLVLIILCSVFAVVTI 300
```


QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDQHFQFPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDQHFQFPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLVMSPEAQNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKDGTAAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKDGTAAL 540
QY 541 SCWADYGAPVFPFLAIGGKDYSEAEALIMTESLNNYPAGDPRLAQAKWEBAFLEM 600
Db 541 SCWADYGAPVFPFLAIGGKDYSEAEALIMTESLNNYPAGDPRLAQAKWEBAFLEM 600
QY 601 RAFQRMAGMFQVTFWAEBSLEDEINRTTAEPLIPATSYIVIFLYISIALGYSWSRV 660
Db 601 RAFQRMAGMFQVTFWAEBSLEDEINRTTAEPLIPATSYIVIFLYISIALGYSWSRV 660
QY 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIFVLE 720
QY 721 YQRLPRRPGEPREVHIGRALGVAPSMMLCSSEALCPFLGALTMPAVRTTALSGLAV 780
Db 721 YQRLPRRPGEPREVHIGRALGVAPSMMLCSSEALCPFLGALTMPAVRTTALSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKQDELPPPGQEGLLIGFFQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKQDELPPPGQEGLLIGFFQKAYAPPL 840
QY 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELPKDSYLLDYFLFLNRYFVGVAPV 900
Db 841 LHMITRGVVLLFLALFGVLSYMSCHISVGLDQELPKDSYLLDYFLFLNRYFVGVAPV 900
QY 901 YFVTTLYGNFSEAGNNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDDFDIW 960
Db 901 YFVTTLYGNFSEAGNNAICSSAGCNPFSTQKIQYATEFPQSYLAIPASSWVDDFDIW 960
QY 961 LTPSSCCRLYISGPNKDFCPTSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPTSVNSLNCNKMSITMGSVRPSVEQPHKYLPMFLNDRP 1020
QY 1021 NIKCPKGGIAAYSTSVNLTSDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGIAAYSTSVNLTSDQVL-----ASRFMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPQTDPAFVFPVPTIINNVEYQYLITLPEGL 1113
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVPQTDPAFVFPVPTIINNVEYQYLITLPEGL 1113
QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSVMLVDTVGFMAWLSYNVAVSLNLYS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSVMLVDTVGFMAWLSYNVAVSLNLYS 1200
QY 1174 AVGMSVEFVSHITRSPAIKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAIKPTWLERAKEATISMGSAVFAGVAMTNLPGLVLGLAKAQ 1260
QY 1234 LQIQIFFRNLNLTLLGLLHGLVFLPVLISYVGPVDPNPALEQKRAEBAVAAMVWASCP 1293
Db 1261 LQIQIFFRNLNLTLLGLLHGLVFLPVLISYVGPVDPNPALEQKRAEBAVAAMVWASCP 1320
QY 1294 NHPRSVSTADNLYVNHSPFGSFKGAGISNPLNNGRQF 1332
Db 1321 NHPRSVSTADNLYVNHSPFGSFKGAGISNPLNNGRQF 1359

RESULT 21
US-10-293-244-1830
; Sequence 1830, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-244-1830

Query Match 99.5%; Score 6872.5; DB 32; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 MABAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
Db 1 MABAGLRGWLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILQKICPRLVTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLILQKICPRLVTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSNQSILFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
Db 121 VNLHCHNTCSNQSILFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPA 180
QY 181 ATLAVGTWGVVGSALCNAQRLNFGDGTGNGLAPLDTTFHLLBPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTWGVVGSALCNAQRLNFGDGTGNGLAPLDTTFHLLBPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQSGDDVATCSQDCAASCPAIPQALDSTFYLGOMPGSLVLIILILCSVFAVVTI 300
Db 241 ARCNEQSGDDVATCSQDCAASCPAIPQALDSTFYLGOMPGSLVLIILILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
Db 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDQHFQFPFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVTELTTPDVELWSPNSQARSEKAFHDQHFQFPFRTNQVILITAPNRSY 420
QY 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLVMSPEAQNISLQDICYAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSLDLDLLELLELQERLRLHQLVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLFYCANAPLTFKDGTAAL 540

Db 481 SLVDCINSLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEEAFLEEM 600
QY 601 RAFORMMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSSWSRV 660
Db 601 RAFORMMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVHIGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
Db 721 YORLPRRPGEPREVHIGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFPKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFPKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLLNRYFEVGAPV 900
Db 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLLNRYFEVGAPV 900
QY 901 YFVTTLGNVFSSEAGNNAICSSAGCNFSFTQKI QYATEFFPQSYLAIPASSWVDDFDW 960
Db 901 YFVTTLGNVFSSEAGNNAICSSAGCNFSFTQKI QYATEFFPQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVNSLCLKNCMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSVNSLCLKNCMSITMGSVRPSVEQFHKLIPWFLNDRP 1020
QY 1021 NIKCPKGGIAAYSTSVNLTSDQVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGIAAYSTSVNLTSDQVLDTVALSPRLEYSCTISAHCNLYLLDSASRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANI TADLRKVPGTDPAFEPFPTIYNVFEQVLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANI TADLRKVPGTDPAFEPFPTIYNVFEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDARSGLNLSIVMILVDTVGFMAWLDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDARSGLNLSIVMILVDTVGFMAWLDISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIPFFRNLITLGLLHGLVFLPVILSYGPDVNPALAEQKAEBAVAAMVASCP 1293
Db 1261 LIQIPFFRNLITLGLLHGLVFLPVILSYGPDVNPALAEQKAEBAVAAMVASCP 1320
QY 1294 NHPRSVSTADNIVNHSFGSITKGAICAI SNFLPNNGRQF 1332
Db 1321 NHPRSVSTADNIVNHSFGSITKGAICAI SNFLPNNGRQF 1359

RESULT 22

US-10-621-758A-44
; Sequence 44, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-621-758A-44

Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPTTIIHQPCYCAFYDECCKNPELSGLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPTTIIHQPCYCAFYDECCKNPELSGLMTLSNVCSLSN 60
QY 61 TPARKITGDHILILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHILILQKICPRLYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNQSIFINVTVAOLGAGOLPAVVAEAFYOHSPAEQSYDSCSRVVPAA 180
Db 121 VNLCHNTCSNQSIFINVTVAOLGAGOLPAVVAEAFYOHSPAEQSYDSCSRVVPAA 180
QY 181 ATLAVGTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGYVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNEQDQDVATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQDQDVATCSQDCAASCPAIPARQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSNQVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAENRSSY 420
Db 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAENRSSY 420
QY 421 RYDLSLLGPKNFSGILDLDLLELQERLRHLQVWSPQARNISLOQICVAPLNPDNT 480
Db 421 RYDLSLLGPKNFSGILDLDLLELQERLRHLQVWSPQARNISLOQICVAPLNPDNT 480
QY 481 SLYDCCINSLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
Db 481 SLYDCCINSLQYFONNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKGTALAL 540
QY 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEEAFLEEM 600
Db 541 SCWADYGAPVFPFLAIGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEEAFLEEM 600
QY 601 RAFORMMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSSWSRV 660
Db 601 RAFORMMAGMFOVTFMAERSLEDEINRTTAEDLPFATSYIVIFLYISIALGSSYSSWSRV 660
QY 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSVGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVSVGADNIFIFVLE 720
QY 721 YORLPRRPGEPREVHIGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
Db 721 YORLPRRPGEPREVHIGRALGRVAPSMMLCSLSEAI CFFLGTALTPMPAVRTFALTSGLA 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFPKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKPQELPPPGQGGELLGFPKAYAPFL 840
QY 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLLNRYFEVGAPV 900
Db 841 LHWITRGVVLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLLNRYFEVGAPV 900

QY 901 YFVTTTLYGNFSSAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTTLYGNFSSAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSTMGSVRPSVEQFHXYLPWFNLDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSTMGSVRPSVEQFHXYLPWFNLDRP 1020
QY 1021 NIKCPKGGAAVSTSVNLTSQGVLT-----ASRFMAVH 1053
DB 1021 NIKCPKGGAAVSTSVNLTSQGVLTVAILSPREYSGTISAHNCNLYLLDSASRFMAVH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGPWALWDISYNAVSLINLVS 1173
DB 1141 FMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGPWALWGISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIPFFRLNLLITLGLLHGLVFLPVILSYVGPVDPVPALAEQKRAEBAVAAMVVASCP 1293
DB 1261 LIQIPFFRLNLLITLGLLHGLVFLPVILSYVGPVDPVPALAEQKRAEBAVAAMVVASCP 1320
QY 1294 NHPRSVSTADNLYVNHSPGSIKAGAI SNFLPNNGRQF 1332
DB 1321 NHPRSVSTADNLYVNHSPGSIKAGAI SNFLPNNGRQF 1359
RESULT 23
US-10-646-301A-44
; Sequence 44, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-646-301A-44
Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLLLOKICRLPYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLLOKICRLPYTGPNTOACCSAKQLVSLASLSITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGOLPAVAYEAFYQHSFAEQSDSCSRVRPAA 180

QY 181 ATLAVGTMCGVYGSALCNAQRWLNFOGDTGNGLAPLDITFHLLEPQOAVSGSIQPLNEGV 240
DB 181 ATLAVGTMCGVYGSALCNAQRWLNFOGDTGNGLAPLDITFHLLEPQOAVSGSIQPLNEGV 240
QY 241 ARCNSQGGDDVATCSODCAASCAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNSQGGDDVATCSODCAASCAPARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGRVAPARDKSMVDPKGTSLSDKLSFSFTHTLTGQFFQCGWGTWASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSMVDPKGTSLSDKLSFSFTHTLTGQFFQCGWGTWASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPFPFRFTNOVILTAPNRSSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHDQHPFPFRFTNOVILTAPNRSSY 420
QY 421 RYDSLILGPNKFSGILLDLDLLELLELQERLRLQVMSPEAQRNLSQDICVAPLNDPT 480
DB 421 RYDSLILGPNKFSGILLDLDLLELLELQERLRLQVMSPEAQRNLSQDICVAPLNDPT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDHFLYCANAPLTFKQGTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPATFATSYIVIFLYISALGSSWSRV 660
DB 601 RAFQRMAGMFOVTTFAERSLEDEINRTTAEDLPATFATSYIVIFLYISALGSSWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVQLQVVPFLVLSGADNIFLVLVE 720
DB 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSILVQLQVVPFLVLSGADNIFLVLVE 720
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPRCQEGELLGFPQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQELPPRCQEGELLGFPQKAYAPFL 840
QY 841 LHWITRGVVLFLFALFGVLSYMSCHISVGLQDEALPKDSYLLDYFLFLNRYFEVGA 900
DB 841 LHWITRGVVLFLFALFGVLSYMSCHISVGLQDEALPKDSYLLDYFLFLNRYFEVGA 900
QY 901 YFVTTTLYGNFSSAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTTLYGNFSSAGNNAICSSAGCNPFSTQKIQYATPEPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSTMGSVRPSVEQFHXYLPWFNLDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVNSLNCNKMSTMGSVRPSVEQFHXYLPWFNLDRP 1020
QY 1021 NIKCPKGGAAVSTSVNLTSQGVLT-----ASRFMAVH 1053
DB 1021 NIKCPKGGAAVSTSVNLTSQGVLTVAILSPREYSGTISAHNCNLYLLDSASRFMAVH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFVFPYTTITNVFYEQVLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGPWALWDISYNAVSLINLVS 1173
DB 1141 FMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGPWALWGISYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGIILVGLAKAQ 1260
QY 1234 LIQIPFFRLNLLITLGLLHGLVFLPVILSYVGPVDPVPALAEQKRAEBAVAAMVVASCP 1293

```
|||||
1261 LIQIFFFRLNLLITLLGLLGLVFLPVILSYVGPDPNPALEQKRAEAAVAVMVASCP 1320
Db
1294 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Qy
1321 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1359
Db
RESULT 24
US-10-663-208A-44
; Sequence 44, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-663-208A-44

Query Match 99.5%; Score 6872.5; DB 36; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

Qy 1 MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFYDECGKRPGLSGSLMTLSNVSCLSN 60
Db 1 MAEAGLRGWLWALLRLLAQSPYTTIHQPGYCAFYDECGKRPGLSGSLMTLSNVSCLSN 60
Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKOLVLEASLSITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTOACCSAKOLVLEASLSITKALLTRCPACSDNF 120
Qy 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
Qy 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQOAVGSGIQPLNEGV 240
Db 181 ATLAVGTMCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQOAVGSGIQPLNEGV 240
Qy 241 ARCNEQGDVATCSQDCQCAAPATARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Db 241 ARCNEQGDVATCSQDCQCAAPATARPQALDSTFYLGQMPGSLVLIILCSVFAVVTI 300
Qy 301 LLVGFVAPARDKSKMVDPKKTSLSKLSFSFTHLLGQFGQGTWVASWPLTILVLSV 360
Db 301 LLVGFVAPARDKSKMVDPKKTSLSKLSFSFTHLLGQFGQGTWVASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTELTDTPVELWSAPNSQARSEKAFHDQHFGPFPFTNQVILTAPNRSY 420
Db 361 IPVVALAAGLVFTELTDTPVELWSAPNSQARSEKAFHDQHFGPFPFTNQVILTAPNRSY 420
Qy 421 RYDLSLLGPNKSGIILDLLELLELQERLRLQVWSPQARNISLQDICVAPLNPDNT 480
Db 421 RYDLSLLGPNKSGIILDLLELLELQERLRLQVWSPQARNISLQDICVAPLNPDNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGFTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGFTALAL 540
```

```
Qy 541 SCMADYGAPVPPFLAIGYKGYKDYSEAEALIMTFTSLNNYPAGDPRLAQAKLWEAFLEBM 600
Db 541 SCMADYGAPVPPFLAIGYKGYKDYSEAEALIMTFTSLNNYPAGDPRLAQAKLWEAFLEBM 600
Qy 601 RAFQRMAGMPQVTTTAAERSLEDEINRTTAAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
Db 601 RAFQRMAGMPQVTTTAAERSLEDEINRTTAAEDLPFATSYIVIFLYISLALGSYSSWSRV 660
Qy 661 MYDSKATILGLGGVAVVLGAVNMAAGFFSYLGIIRSLVILQVVPFLVLSVGDNIPIFVLE 720
Db 661 MYDSKATILGLGGVAVVLGAVNMAAGFFSYLGIIRSLVILQVVPFLVLSVGDNIPIFVLE 720
Qy 721 YORLPRRPGEPREVIHIGRALGRVAPSMCLLSLSAICFFLGAITPMPAVRFAITSLGLAV 780
Db 721 YORLPRRPGEPREVIHIGRALGRVAPSMCLLSLSAICFFLGAITPMPAVRFAITSLGLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVXPQELPPPPQGGEGLLGFPQKAYAPPL 840
Db 781 ILDFLLQMSAFVALLSLDSKQEAERLDVCCVXPQELPPPPQGGEGLLGFPQKAYAPPL 840
Qy 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPVGPAPV 900
Db 841 LHWITRGVLLFLALFGVLSYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPVGPAPV 900
Qy 901 YFVITLGYNPSSEAGMNAICSSAGCNPFSFTQKIQYATEFFPEQSYLAIPASSWVDDFDW 960
Db 901 YFVITLGYNPSSEAGMNAICSSAGCNPFSFTQKIQYATEFFPEQSYLAIPASSWVDDFDW 960
Qy 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMCSITMGSVRPSVEQPHKYLWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKMCSITMGSVRPSVEQPHKYLWFLNDRP 1020
Qy 1021 NIKCPKGLAAYSTVNLTSQGVL-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAYSTVNLTSQGVLDTVAIILSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSQDYTBALRAARELAANITADLRKVPCTDPAPVFPVYITNVPVEQYLTILPEGL 1113
Db 1054 KPLKNSQDYTBALRAARELAANITADLRKVPCTDPAPVFPVYITNVPVEQYLTILPEGL 1113
Qy 1114 FMLSCLVPTTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGPALMDISYNAVSLINLVS 1173
Db 1114 FMLSCLVPTTFAVSCLLGLDLRLSGLNLLSIVMLVDTVGPALMDISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAVAGVAMTNLPGLVILGLAKAQ 1233
Db 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISGSAVFAVAGVAMTNLPGLVILGLAKAQ 1260
Qy 1234 LIQIFFFRLNLLITLLGLLGLVFLPVILSYVGPDPNPALEQKRAEAAVAVMVASCP 1293
Db 1234 LIQIFFFRLNLLITLLGLLGLVFLPVILSYVGPDPNPALEQKRAEAAVAVMVASCP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1332
Db 1294 NHPSRVSTADNIYVNHSPFGSIKGAGAI SNFLPNNGRQF 1359
```

```
RESULT 25
US-10-736-769-44
; Sequence 44, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
```

; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-736-769-44

Query Match 99.5%; Score 6872.5; DB 37; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLTSLNSVCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLTSLNSVCLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGQLPAVVAEYFQHSFABQSYDSCSRVRPAA 180
DB 121 VNLHCHNTCSPNOSLFINTVTRVAQLGAGQLPAVVAEYFQHSFABQSYDSCSRVRPAA 180

QY 181 ATLAUVTGCVYGSALCNARWLNFGDGTGNGLAPLDITPHLLPEQAVGSGIQLNEGV 240
DB 181 ATLAUVTGCVYGSALCNARWLNFGDGTGNGLAPLDITPHLLPEQAVGSGIQLNEGV 240

QY 241 ARCNEQSGDDVATCSQDCAASCPAIAARPOALDSTFYLQMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNEQSGDDVATCSQDCAASCPAIAARPOALDSTFYLQMPGSLVLIILCSVPAVVTI 300

QY 301 LVLGFRVAPARDKSKMVDPKGTSLSDKLSFSHTLLGQFFQCGMTWVASWPLTILVLSV 360
DB 301 LVLGFRVAPARDKSKMVDPKGTSLSDKLSFSHTLLGQFFQCGMTWVASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPERSSY 420
DB 361 IPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAFHDQHPGFFRTNQVILTAPERSSY 420

QY 421 RYDSSLGPKNPSGIIIDLLLELLEQLERLRLQVWSPQAORNISLQDICVAPLNPNT 480
DB 421 RYDSSLGPKNPSGIIIDLLLELLEQLERLRLQVWSPQAORNISLQDICVAPLNPNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLALTANQTLMGQTSQVDWKDHFYCANAPLTKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLALTANQTLMGQTSQVDWKDHFYCANAPLTKDGTALAL 540

QY 541 SCHADYGAPVFPFLAIGYKGKDYSAEALIMTFPSLNYPAGDPRLAQAKWEEAPLEM 600
DB 541 SCHADYGAPVFPFLAIGYKGKDYSAEALIMTFPSLNYPAGDPRLAQAKWEEAPLEM 600

QY 601 RAFORMAGMFWTFAERSLDEINRTTAEDLPFATSYIVIFLYISALGYSYSWSRV 660
DB 601 RAFORMAGMFWTFAERSLDEINRTTAEDLPFATSYIVIFLYISALGYSYSWSRV 660

QY 661 MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720
DB 661 MVDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIFIVLE 720

QY 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTFALTSLGLAV 780
DB 721 YQRLPRPGEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTFALTSLGLAV 780

QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPGCGEGILLGPFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDVCCCKVQOELPPGCGEGILLGPFQKAYAPFL 840

QY 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900
DB 841 LHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGA 900

QY 901 YFVTTILGYNPFSSSEAGNNAICSSAGCNPFSTQKIQVATEPPEQSILAIAPASSWDDPIDW 960
DB 901 YFVTTILGYNPFSSSEAGNNAICSSAGCNPFSTQKIQVATEPPEQSILAIAPASSWDDPIDW 960

QY 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRSPSVQFHKYLPWFNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCILKNCMSITMGSVRSPSVQFHKYLPWFNDRP 1020

QY 1021 NIKCPKGLAAAYSTSVNLTSDDQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGLAAAYSTSVNLTSDDQVLDTVAILSPREYSGTISAHCNLYLLDSASRFMAYH 1080

QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFVPPYTTITNVFYEQYLTILPEGL 1113
DB 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFVPPYTTITNVFYEQYLTILPEGL 1113

QY 1114 FMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLVS 1173
DB 1141 FMLSCLVPTFAVSCILLGLDLRSGLLNLISIVMILVDTVGFMAWDISYNAVSLINLVS 1200

QY 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAVAMTNLPGILVLGLAKAQ 1260

QY 1234 LIQIIPFRLNLIITLGLLHGLVFLPVILSYGPDVNPALALBOKRAEBAVAAVMVASCP 1293
DB 1261 LIQIIPFRLNLIITLGLLHGLVFLPVILSYGPDVNPALALBOKRAEBAVAAVMVASCP 1320

QY 1294 NHPSRVSTADNIYVNHSPFGSFKGAGAINFLPNNGRQF 1332
DB 1321 NHPSRVSTADNIYVNHSPFGSFKGAGAINFLPNNGRQF 1359

RESULT 26
US-10-750-386-44
; Sequence 44, Application US/10750386
; GENERAL INFORMATION:
; APPLICANT: Garcia-Calvo, Margarita
; APPLICANT: Chapman, Kevin
; APPLICANT: Goulet, Mark
; APPLICANT: Ujainwalla, Peroze
; APPLICANT: Altmann, Scott W
; APPLICANT: Davis, Chip
; APPLICANT: Bull, Herb
; APPLICANT: Thornberry, Nancy A
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING AGONISTS AND ANTAGONIST
; FILE REFERENCE: A36104 074669 0103
; CURRENT APPLICATION NUMBER: US/10/750,386
; CURRENT FILING DATE: 2003-12-31
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 44
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-750-386-44

Query Match 99.5%; Score 6872.5; DB 37; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLTSLNSVCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLTSLNSVCLSN 60

QY 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVSLSEASITKALLTRCPACSDNF 120

QY 121 VNLCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
DB 121 VNLCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAGVTMCGVYGSALCNARWLNFQDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAGVTMCGVYGSALCNARWLNFQDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDVATCSQDCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNSQGGDVATCSQDCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSNVDPKKGTSLSKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVRVAPARDKSNVDPKKGTSLSKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILTAPNRSY 420
DB 361 IPVVALAAGLVTELTTPDVELWSAPNSQARSEKAFHQHFGPFRTNQVILTAPNRSY 420
QY 421 RYDSLLGPKNPSGILDLDDLELLELOBLRHLQVMSPEAQRNISLODICYAPLNPONT 480
DB 421 RYDSLLGPKNPSGILDLDDLELLELOBLRHLQVMSPEAQRNISLODICYAPLNPONT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDWKDHELYCANAPLTFKDGTTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDWKDHELYCANAPLTFKDGTTALAL 540
QY 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAKWEAEAFLEEM 600
QY 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIIVIFLYISALGYSYWSRV 660
DB 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIIVIFLYISALGYSYWSRV 660
QY 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIIFFVLE 720
DB 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSLVILQVVPFLVLSVGADNIIFFVLE 720
QY 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CPFELGALTMPAVRTFALTSGLAV 780
DB 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLSEAI CPFELGALTMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKROEASRLDYCCCKVQOELPPQCGEGILLGPFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKROEASRLDYCCCKVQOELPPQCGEGILLGPFQKAYAPFL 840
QY 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHI SVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTLLGYNFSSEAGMNAICSSAGCNFFSTQKIQVATEFPQSYLAIPASSWVDFFIDW 960
DB 901 YFVTTLLGYNFSSEAGMNAICSSAGCNFFSTQKIQVATEFPQSYLAIPASSWVDFFIDW 960
QY 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCIKNCSITMGSVRPSVEQFKHYLPWFANDRP 1020
DB 961 LTPSSCCRLYISGPNKDKCPSTVNSLNCIKNCSITMGSVRPSVEQFKHYLPWFANDRP 1020
QY 1021 NIKCPKGGLLAAYSTSVNLSYSDGQVL-----ASRPMAYH 1053
DB 1021 NIKCPKGGLLAAYSTSVNLSYSDGQVL-----ASRPMAYH 1053
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPDPAFEVFPVYITNMFVEQYLTILPEGL 1113
DB 1054 KPLKNSQDYTEALRAARELAANITADLRKVPDPAFEVFPVYITNMFVEQYLTILPEGL 1113
QY 1114 FMLSCLVPTFAVSCLLGLDLSGLNLLSIIMVILVDTGFMALWDISYNAVSLINLYS 1173
DB 1141 FMLSCLVPTFAVSCLLGLDLSGLNLLSIIMVILVDTGFMALWDISYNAVSLINLYS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSVAVFAGVAMTNLPGILVLGLAKAQ 1233

DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSVAVFAGVAMTNLPGILVLGLAKAQ 1260
QY 1234 LIQIFFFFRLNLLITLLGLLHGLVFLPVLTLSVVGDPVNPALALEQKRAEEAAVAVNVASCP 1293
DB 1261 LIQIFFFFRLNLLITLLGLLHGLVFLPVLTLSVVGDPVNPALALEQKRAEEAAVAVNVASCP 1320
QY 1294 NHPRSVSTADNIYVNHSPSGIKGAGAISNPLPNNGRQP 1332
DB 1321 NHPRSVSTADNIYVNHSPSGIKGAGAISNPLPNNGRQP 1359

RESULT 27

US-11-218-141-1830
; Sequence 1830, Application US/11218141
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/11/218,141
; CURRENT FILING DATE: 2005-08-31
; PRIOR APPLICATION NUMBER: Not yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 1830
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-218-141-1830

Query Match 99.5%; Score 6872.5; DB 42; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;
Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;
QY 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVSCLSN 60
QY 61 TPARKITGDHLILLLOKICPRLVTGNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
DB 61 TPARKITGDHLILLLOKICPRLVTGNTQACSAKOLVLSLEASLSITKALLTRCPACSDNF 120
QY 121 VNLCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
DB 121 VNLCHNTCSNQSIFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDSCSRVRPAA 180
QY 181 ATLAGVTMCGVYGSALCNARWLNFQDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
DB 181 ATLAGVTMCGVYGSALCNARWLNFQDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV 240
QY 241 ARCNSQGGDVATCSQDCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
DB 241 ARCNSQGGDVATCSQDCAASCPAARPOALDSTFYLGOMPGSLVLIILCSVPAVVTI 300
QY 301 LLVGFVRVAPARDKSNVDPKKGTSLSKLSFSTHTLLGQFQCGWGTWVASWPLTILVLSV 360


```
Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKUSFSFTHLLGQFPQGWGTWASWPLTILVLSV 360
Qy 361 IPVVALAAGLVFTTELTPDPVELWSAPNSQARSEKAFHDQHFQGFPRFTNOVILTAPNRSSV 420
Db 361 IPVVALAAGLVFTTELTPDPVELWSAPNSQARSEKAFHDQHFQGFPRFTNOVILTAPNRSSV 420
Qy 421 RYDSLLGPKNPGSGLDLDLLELLELQERLRLHQLWSPEAQNISLQDICYAPLNPNT 480
Db 421 RYDSLLGPKNPGSGLDLDLLELLELQERLRLHQLWSPEAQNISLQDICYAPLNPNT 480
Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLXCANAPLTKDGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLXCANAPLTKDGTALAL 540
Qy 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKLEAEAPLEEM 600
Db 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKLEAEAPLEEM 600
Qy 601 RAPQRMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Db 601 RAPQRMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660
Qy 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLVQVVPPIVLSYGADNIFIFVLE 720
Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLVQVVPPIVLSYGADNIFIFVLE 720
Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGSLAV 780
Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGSLAV 780
Qy 781 ILDFLLQMSAFVALLSLDSKRQBSRLDVCCVKQELPPPGQEGELGFGFKAYAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQBSRLDVCCVKQELPPPGQEGELGFGFKAYAPFL 840
Qy 841 LHWITRGVLLPLALPGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPV 900
Db 841 LHWITRGVLLPLALPGVLSYMSCHISVGLDQELALPKDSYLLDYFLFNRYFEVGAPV 900
Qy 901 YFVTTLYGNFSSBAGNNAICSSAGCNPFSTOKIOYATPPEQSYLAIPASSWVDDFDIM 960
Db 901 YFVTTLYGNFSSBAGNNAICSSAGCNPFSTOKIOYATPPEQSYLAIPASSWVDDFDIM 960
Qy 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMCSITMGSVRPSVEQFHKLIPWFLNDRP 1020
Db 961 LTPSSCCRLYISGPNKDFCPSTVNSLNCNKMCSITMGSVRPSVEQFHKLIPWFLNDRP 1020
Qy 1021 NIKCPKGLAAVSTSNLTSQGVLT-----ASRPMAYH 1053
Db 1021 NIKCPKGLAAVSTSNLTSQGVLTVAILSPRLEYSGTISAHCNLYLLDSASRPMAYH 1080
Qy 1054 KPLKNSODYTEALRAARELAANITADLRKVGCTDPAFEPVPTITNVFYEQYLTILPEGL 1113
Db 1081 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFEPVPTITNVFYEQYLTILPEGL 1140
Qy 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWDISYNAVSLINLVS 1173
Db 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDVTGFMALWDISYNAVSLINLVS 1200
Qy 1174 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVPAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LQIQIFFRNLTLTLGLLHGLVFLPVILSYGPDVNPALAEQKAEBAVAAMVASCP 1293
Db 1261 LQIQIFFRNLTLTLGLLHGLVFLPVILSYGPDVNPALAEQKAEBAVAAMVASCP 1320
Qy 1294 NHPRSVSTADNTVYVNSHSPGSIKGAGAI SNFLPNNGROF 1332
Db 1321 NHPRSVSTADNTVYVNSHSPGSIKGAGAI SNFLPNNGROF 1359
```

```
US-10-170-205E-21686
; Sequence 21686, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; TITLE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CL001381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 21686
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-21686
```

Query Match 99.4%; Score 6868.5; DB 31; Length 1359;

Best Local Similarity 97.8%; Pred. No. 0;

Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

Qy 1 MABAGLRGWLWALLRLLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCLSN 60

Db 1 MABAGLRGWLWALLRLLAQSEPYTTIHOPGYCAFYDECGKNPELSGSLMTLSNVCLSN 60

Qy 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120

Db 61 TPARKITGDHLILLQKICPRLYTGPNTOACCSAKQLVSLSEASISITKALLTRCPACSDNF 120

Qy 121 VNLHCHTCSNQSLFINVTRVAQAGAGQLPAVAYEAFYQHSFASQSYDSCSRVRPAA 180

Db 121 VNLHCHTCSNQSLFINVTRVAQAGAGQLPAVAYEAFYQHSFASQSYDSCSRVRPAA 180

Qy 181 ATLAVGTWGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240

Db 181 ATLAVGTWGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQLNEGV 240

Qy 241 ARCNSQGDVATCSQDCCAACPAIARPOALDSTFYLQMPGSLVLIILCSVPAVVTI 300

Db 241 ARCNSQGDVATCSQDCCAACPAIARPOALDSTFYLQMPGSLVLIILCSVPAVVTI 300

Qy 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360

Db 301 LLVGRVAPARDKSKWDPKKGTSLSDKLSFSTHTLLGQFPQGWGTWASWPLTILVLSV 360

Qy 361 IPVVALAAGLVFTTELTPDPVELWSAPNSQARSEKAFHDQHFQGFPRFTNOVILTAPNRSSV 420

Db 361 IPVVALAAGLVFTTELTPDPVELWSAPNSQARSEKAFHDQHFQGFPRFTNOVILTAPNRSSV 420

Qy 421 RYDSLLGPKNPGSGLDLDLLELLELQERLRLHQLWSPEAQNISLQDICYAPLNPNT 480

Db 421 RYDSLLGPKNPGSGLDLDLLELLELQERLRLHQLWSPEAQNISLQDICYAPLNPNT 480

Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLXCANAPLTKDGTALAL 540

Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHLXCANAPLTKDGTALAL 540

Qy 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKLEAEAPLEEM 600

Db 541 SCMDYGAIVFPFLAIGGYKGDYSEAEALIMTFSLNYPAGDPRLAQAQKLEAEAPLEEM 600

Qy 601 RAPQRMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660

Db 601 RAPQRMAGMFOVTTFAERSLEDEINRTTAEDLPIPATSYIVIFLYISIALGYSYSSWSRV 660

Qy 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLVQVVPPIVLSYGADNIFIFVLE 720

Db 661 MYDSKATLGLGGVAVVLGAVMAAMGFFSYLGRSSVLVQVVPPIVLSYGADNIFIFVLE 720

Qy 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGSLAV 780

Db 721 YQRLPRRPGEPREVIHGRALGRVAPSMMLCSLSEAI CFFLGALTMPAVRTFALTSGSLAV 780

```

QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQOELPPPPQOGBGLLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQOELPPPPQOGBGLLLGFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTGLYNFSSAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTGLYNFSSAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNCLNCSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNCLNCSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDDQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGGLAAYSTSVNLTSDDQVLDTVAILSPRLEYSCTISAHCNLYLLDSTSRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFYEQYLTILPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFYEQYLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTVGFMAWGLSYNAVSLINLVS 1173
DB 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTVGFMAWGLSYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1260
QY 1234 LIQIPFFRLNLLITLGLLHGLVFLPVLISYVGPDPVPALAEQKRAEBAVAANVASC 1293
DB 1261 LIQIPFFRLNLLITLGLLHGLVFLPVLISYVGPDPVPALAEQKRAEBAVAANVASC 1320
QY 1294 NHPRSVSTADNLYVNHSPGSGTKGAGISNLPNNGRQF 1332
DB 1321 NHPRSVSTADNLYVNHSPGSGTKGAGISNLPNNGRQF 1359

RESULT 29
US-60-453-050-8956
; Sequence 8956, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001457
; CURRENT APPLICATION NUMBER: US/60/453,050
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-050-8956

Query Match 99.4%; Score 6868.5; DB 48; Length 1359;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

QY 1 MAEAGLRGMLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
DB 1 MAEAGLRGMLLWALLRLAQSEPYTHIQPGYCAFYDECGKNPELSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLILQKICRPLYTGPNTOACCSAKQLVLEASLSITKALLITRCPACSDNF 120
DB 61 TPARKITGDHLLILQKICRPLYTGPNTOACCSAKQLVLEASLSITKALLITRCPACSDNF 120
QY 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180

```

```

DB 121 VNLHCHNTCPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSDSCSRVRPAA 180
QY 181 ATLAGVTMCYGSALCNAQRWLNFOGDTGNGLAPLIDITFHLLEPGQAVGSGIQLPNEG 240
DB 181 ATLAGVTMCYGSALCNAQRWLNFOGDTGNGLAPLIDITFHLLEPGQAVGSGIQLPNEG 240
QY 241 ARCHNESQDDVATCSCODCAASCPAIARPOALDSTFVLGQMPGSLVLIIILCSVFVV 300
DB 241 ARCHNESQDDVATCSCODCAASCPAIARPOALDSTFVLGQMPGSLVLIIILCSVFVV 300
QY 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVWASWPLTIL 360
DB 301 LLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGMGTWVWASWPLTIL 360
QY 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGPPRTNQVILITAPNRS 420
DB 361 IPVVALAAGLVFTLTTPVELWSAPNSQARSEKAFHDOHFGPPRTNQVILITAPNRS 420
QY 421 RYDSLILGPKNFGSLDLDLLELLELQERLRLHQVMSPEAQNISLQDICYAPLNPDNT 480
DB 421 RYDSLILGPKNFGSLDLDLLELLELQERLRLHQVMSPEAQNISLQDICYAPLNPDNT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDMKOHFLYCANNAPLTFKOGTAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANOTLMGOTSQVDMKOHFLYCANNAPLTFKOGTAL 540
QY 541 SCMADYGAPVFPFLAIGYKGYKDYSEALITMTFSLANNYPAGDPRLAQAKLWEEAFLE 600
DB 541 SCMADYGAPVFPFLAIGYKGYKDYSEALITMTFSLANNYPAGDPRLAQAKLWEEAFLE 600
QY 601 RAFORMMAGMVOVTFMAERSLEDEINRTTADLPFATSYIVIFLYISIALGYSWSSRV 660
DB 601 RAFORMMAGMVOVTFMAERSLEDEINRTTADLPFATSYIVIFLYISIALGYSWSSRV 660
QY 661 MVDSKATILGLGGVAVVLGAVMAAMGFFSYLGIRSLVILQVVPFVLVSVGADNIFIVLE 720
DB 661 MVDSKATILGLGGVAVVLGAVMAAMGFFSYLGIRSLVILQVVPFVLVSVGADNIFIVLE 720
QY 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSISEAICPFLGALTMPAVRTFALTSGLAV 780
DB 721 YORLPRRPGEPREVIHGRALGRVAPSMILCSISEAICPFLGALTMPAVRTFALTSGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQOELPPPPQOGBGLLLGFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQESRLDVCCCKVQOELPPPPQOGBGLLLGFQKAYAPFL 840
QY 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
QY 901 YFVTTGLYNFSSAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
DB 901 YFVTTGLYNFSSAGMNAICSSAGCNFSFTQKIQYATEFPEQSYLAIPASSWVDDFDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNCLNCSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLCLNCLNCSITMGSVRPSVEQFHKYLPLWFLNDRP 1020
QY 1021 NIKCPKGGLAAYSTSVNLTSDDQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGGLAAYSTSVNLTSDDQVLDTVAILSPRLEYSCTISAHCNLYLLDSTSRFMAYH 1080
QY 1054 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFYEQYLTILPEGL 1113
DB 1081 KPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTITNVFYEQYLTILPEGL 1140
QY 1114 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTVGFMAWGLSYNAVSLINLVS 1173
DB 1141 FMLSCLVPTFAVSCLLGLDLRSGLNLLSVMLIVDTVGFMAWGLSYNAVSLINLVS 1200
QY 1174 AVGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGLILVLGLAKAQ 1233

```

Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260
Qy 1234 LIQIFPRLNLITLLGLHGLVFLPVILSYGPDVNPALALBQKRAEBAVAAVMVASCP 1293
Db 1261 LIQIFPRLNLITLLGLHGLVFLPVILSYGPDVNPALALBQKRAEBAVAAVMVASCP 1320
Qy 1294 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1359

RESULT 30
US-60-453-135-8956
; Sequence 8956, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT APPLICATION NUMBER: US/60/453,135
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-453-135-8956

Query Match 99.4%; Score 6868.5; DB 48; Length 1359;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 1329; Conservative 0; Mismatches 3; Indels 27; Gaps 1;

Qy 1 MABAGURGLLWALLRLAQSEPYTHOPGYCAFYDECGKPELGSGLMTLSNVCLSN 60
Db 1 MABAGURGLLWALLRLAQSEPYTHOPGYCAFYDECGKPELGSGLMTLSNVCLSN 60

Qy 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLITRCPACSDNF 120
Db 61 TPARKITGDHLLILLOKICPRLYTGPNTQACCSAKQLVLSLEASISITKALLITRCPACSDNF 120

Qy 121 VNLHCHTCSNQSLEFINTVRAQLGAGQLPAVAYEAFYQHSFABQSDYSCSRVRPAA 180
Db 121 VNLHCHTCSNQSLEFINTVRAQLGAGQLPAVAYEAFYQHSFABQSDYSCSRVRPAA 180

Qy 181 ATLAAGTMCVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLERQQAQVSGIQPLNEGY 240
Db 181 ATLAAGTMCVGSALCNAQRWLNFGDGTGNGLAPLDITPHLLERQQAQVSGIQPLNEGY 240

Qy 241 ARCNESQGDVATCSQDCDCAASCPAIPARQALDSTFYLQMPGSLVLIILICSVPAVVTI 300
Db 241 ARCNESQGDVATCSQDCDCAASCPAIPARQALDSTFYLQMPGSLVLIILICSVPAVVTI 300

Qy 301 LLVGFVAPARKSRVDPKGTSLSDKLSPSTHTLLGQFPQGWGTWVASWPLTILVLSV 360
Db 301 LLVGFVAPARKSRVDPKGTSLSDKLSPSTHTLLGQFPQGWGTWVASWPLTILVLSV 360

Qy 361 IPVVALAAGLVTELTTPDVELWSAQNQAARSEKAFHDQHPFFRTNQVILITAPNRSY 420
Db 361 IPVVALAAGLVTELTTPDVELWSAQNQAARSEKAFHDQHPFFRTNQVILITAPNRSY 420

Qy 421 RYDLSLLGPKNFGSILDLLLELLELQERLHLQVWSPQAORNSLQDICVAPLNPDNT 480
Db 421 RYDLSLLGPKNFGSILDLLLELLELQERLHLQVWSPQAORNSLQDICVAPLNPDNT 480

Qy 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKQGTALAL 540
Db 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKQGTALAL 540

Qy 541 SCHADYGAPVFFFLAIGGKGYKDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLEEM 600
Db 541 SCHADYGAPVFFFLAIGGKGYKDYSEABALIMTFSLNNTYPAGDPRLAQAKLWEEAFLEEM 600

Qy 601 RAFQRMAGMFQVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660
Db 601 RAFQRMAGMFQVTFATERSLEDEINRTTAEDLPIFATSYIVIFLYISLALGSYSMSRV 660

Qy 661 MVDKATLGLGGVAVVILGAVMAAGFFSYLIGRISLVLQVVPFLVLSVGDADNIFIPVLE 720
Db 661 MVDKATLGLGGVAVVILGAVMAAGFFSYLIGRISLVLQVVPFLVLSVGDADNIFIPVLE 720

Qy 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLEAICFFLGALTTPMPAVRTEALTSLAV 780
Db 721 YQRLPRRPGEPREHIGRALGRVAPSMLLCSLEAICFFLGALTTPMPAVRTEALTSLAV 780

Qy 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCKVQELPPPGQEGGLLGFQKAVAPFL 840
Db 781 ILDFLLQMSAFVALLSLDSKRQASRLDVCCKVQELPPPGQEGGLLGFQKAVAPFL 840

Qy 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYLFLNRYFEVAPV 900
Db 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYLFLNRYFEVAPV 900

Qy 901 YFVTTILGYNFSSSEAGNNAICSSAGCANNFSTQKIQYATBPEOSYLAIPASSWVDDFIDW 960
Db 901 YFVTTILGYNFSSSEAGNNAICSSAGCANNFSTQKIQYATBPEOSYLAIPASSWVDDFIDW 960

Qy 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNLCNCSITMGSVRPSVQFHKYLPWFNLDRP 1020
Db 961 LTPSSCCRLYISGPNKDKFCPTVNSLNCNLCNCSITMGSVRPSVQFHKYLPWFNLDRP 1020

Qy 1021 NIKCPKGGLAAYSTSVNLTSQGVL-----ASRFMAYH 1053
Db 1021 NIKCPKGGLAAYSTSVNLTSQGVL-----ASRFMAYH 1053

Qy 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFVPPYTTITNVFYEQVLTILPEGL 1113
Db 1054 KPLKNSQDYTEALRAARELAANITADLRKVGCTDPAFVPPYTTITNVFYEQVLTILPEGL 1113

Qy 1114 FMLSCLVPTFAVSCILLGLSLNLLSVMLILVDTVGFMAWDISYNAVSLNLVLS 1173
Db 1141 FMLSCLVPTFAVSCILLGLSLNLLSVMLILVDTVGFMAWDISYNAVSLNLVLS 1200

Qy 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1233
Db 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQ 1260

Qy 1234 LIQIFPRLNLITLLGLHGLVFLPVILSYGPDVNPALALBQKRAEBAVAAVMVASCP 1293
Db 1261 LIQIFPRLNLITLLGLHGLVFLPVILSYGPDVNPALALBQKRAEBAVAAVMVASCP 1320

Qy 1294 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1332
Db 1321 NHPSRVSTADNIYVNHSPFGSIKAGAINSNFLPNNGRQF 1359

RESULT 31
US-60-466-412-8956
; Sequence 8956, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8956
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-8956

Query Match		99.4%;	Score 6868.5;	DB 48;	Length 1359;	
Best Local Similarity		97.8%;	Pred. No. 0;			
Matches 1329;		Conservative	0;	Mismatches	3;	Indels 27; Gaps 1;
Qy	1	MAEAGLGRWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKPELGSGLMTLSNVSCLSN	60			
Db	1	MAEAGLGRWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKPELGSGLMTLSNVSCLSN	60			
Qy	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Db	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Qy	121	VNLHCHNTCSPNQSIFINVTVAQIGAGOLPAVVAEAFYQHSFAEQSDSCSRVPPAA	180			
Db	121	VNLHCHNTCSPNQSIFINVTVAQIGAGOLPAVVAEAFYQHSFAEQSDSCSRVPPAA	180			
Qy	181	ATLAVGTMCVGVGSAALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240			
Db	181	ATLAVGTMCVGVGSAALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240			
Qy	241	ARCNSQGDVATCSDQDCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300			
Db	241	ARCNSQGDVATCSDQDCAASCPAIRPQALDSTFYLGQMPGSLVLIILCSVFAVVTI	300			
Qy	301	LLVGERVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQGWGTTWASWPLTILVLSV	360			
Db	301	LLVGERVAPARDKSKWDPKGTSLSDKLSFSTHTLLGQFQGWGTTWASWPLTILVLSV	360			
Qy	361	IPVALLAAGLVTELTTPDVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTAENRSSY	420			
Db	361	IPVALLAAGLVTELTTPDVELWSAPNSQARSEKAFHQHFGFPFRFTNQVILTAENRSSY	420			
Qy	421	RYDSLLGLPKNSGILDLLELLELQERLRLHQLVMSPEAQRNLSLDICVAPLNPONT	480			
Db	421	RYDSLLGLPKNSGILDLLELLELQERLRLHQLVMSPEAQRNLSLDICVAPLNPONT	480			
Qy	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKDGTTALAL	540			
Db	481	SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVMDKDFLYCANAPLTPKDGTTALAL	540			
Qy	541	SCWADYGAPVFFFLAIGGKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM	600			
Db	541	SCWADYGAPVFFFLAIGGKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEEAFLEEM	600			
Qy	601	RAFQRMAGMFQVTFATRSLSDEINRTTAEDLPIFATSYIIVIFLYISALGSYSWSRV	660			
Db	601	RAFQRMAGMFQVTFATRSLSDEINRTTAEDLPIFATSYIIVIFLYISALGSYSWSRV	660			
Qy	661	MVDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSGADNIFIFVLE	720			
Db	661	MVDSKATLGLGGVAVVLGAVMAAGFFSYLGIRSSLVILQVVPFLVLSGADNIFIFVLE	720			
Qy	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEALICFFLGALTPMPAVRTTALTSLGLAV	780			
Db	721	YQRLPRRPGEPREHVHIGRALGRVAPSMLLCSLSEALICFFLGALTPMPAVRTTALTSLGLAV	780			
Qy	781	ILDFLLQMSAFVALLSLSKQRBASRLDVCCCKVQPELPPPGQCEGLLIGFTQKAYAPFL	840			
Db	781	ILDFLLQMSAFVALLSLSKQRBASRLDVCCCKVQPELPPPGQCEGLLIGFTQKAYAPFL	840			
Qy	841	LHWITRGVLLLFALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900			
Db	841	LHWITRGVLLLFALFGVSLYSWCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV	900			
Qy	901	YFVTTILGYNFSSEAGMNAICSSAGCNFSFTQIKIYATEPPEQSYLAIPASSWVDFFIDW	960			
Db	901	YFVTTILGYNFSSEAGMNAICSSAGCNFSFTQIKIYATEPPEQSYLAIPASSWVDFFIDW	960			
Qy	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSTMGSVRPSVFQPHKYLFWFLNDRP	1020			
Db	961	LTPSSCCRLYISGPNKDKFCPSTVNSLNCNKCMSTMGSVRPSVFQPHKYLFWFLNDRP	1020			
Qy	1021	NIKCPKGLAAYSTSVNLTSDGQVLT-----ASRFMAYH	1053			

Db	1021	NIKCPKGLAAYSTSVNLTSDGQVLT-----ASRFMAYH	1080			
Qy	1054	KPLKNSQDYTEALPAARELANITADLRKVPCTDPAPFVPPYTTITNNVFEYQYLTLPEGL	1113			
Db	1081	KPLKNSQDYTEALPAARELANITADLRKVPCTDPAPFVPPYTTITNNVFEYQYLTLPEGL	1140			
Qy	1114	FMLSICLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS	1173			
Db	1141	FMLSICLVPTFAVSCLLGLDLRSGLNLLSIVMLVDTVGFMALWDISYNAVSLINLVS	1200			
Qy	1174	AVGMSVFEVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1233			
Db	1201	AVGMSVFEVSHITRSFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ	1260			
Qy	1234	LIQIPFFRLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEQKRAEBAEAAVAVASCP	1293			
Db	1261	LIQIPFFRLNLLITLLGLLHGLVFLPVILSVGPDVNPALALEQKRAEBAEAAVAVASCP	1320			
Qy	1294	NHPSRVSTADNIYVNHSPFGSISKGAGAINSLFNPNGROF	1332			
Db	1321	NHPSRVSTADNIYVNHSPFGSISKGAGAINSLFNPNGROF	1359			
RESULT 32						
PCT-US01-08631-53052						
; Sequence 53052, Application PC/TUS0108631						
; GENERAL INFORMATION:						
; APPLICANT: Hyseq, Inc						
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES						
; FILE REFERENCE: 21272-049						
; CURRENT APPLICATION NUMBER: PCT/US01/08631						
; CURRENT FILING DATE: 2001-03-30						
; PRIOR FILING DATE: 09/540,217						
; PRIOR FILING DATE: 2000-03-31						
; PRIOR APPLICATION NUMBER: 09/649,167						
; PRIOR FILING DATE: 2000-08-23						
; NUMBER OF SEQ ID NOS: 60736						
; SOFTWARE: Custom						
; SEQ ID NO 53052						
; LENGTH: 1344						
; TYPE: PRT						
; ORGANISM: Homo sapiens						
; FEATURE:						
; NAME/KEY: DOMAIN						
; LOCATION: (901)..(917)						
; OTHER INFORMATION: OBSTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,						
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87						
; NAME/KEY: misc.feature						
; LOCATION: (1)..(1344)						
; OTHER INFORMATION: Xaa = X or * as defined in Table 2						
PCT-US01-08631-53052						
Query Match						
Best Local Similarity 94.6%; Score 6536; DB 1; Length 1344;						
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;						
Qy	1	MAEAGLGRWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKPELGSGLMTLSNVSCLSN	60			
Db	1	MAEAGLGRWLLWALLRLRLAQSEPYTHIQPGYCAFYDECGKPELGSGLMTLSNVSCLSN	60			
Qy	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Db	61	TPARKITGDHLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNF	120			
Qy	121	VNLHCHNTCSPNQSIFINVTVAQIGAGOLPAVVAEAFYQHSFAEQSDSCSRVPPAA	180			
Db	121	VNLHCHNTCSPNQSIFINVTVAQIGAGOLPAVVAEAFYQHSFAEQSDSCSRVPPAA	180			
Qy	181	ATLAVGTMCVGVGSAALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240			
Db	181	ATLAVGTMCVGVGSAALCNAQRWLNFGDGTGNGLAPLDITFHLLPEQAVGSGIQPLNEGV	240			

```
QY 241 ARCNSQDDVATCSQDCAAACP--AIARPOALDSTFYLGQMPGSLVLIILCSVFVV 298
Db 241 ARCNSQDDVATCSQDCAAACPXPAPRPSTPSTW--RQMPGSLVLIILCSVFVV 298
QY 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVL 358
Db 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVL 358
QY 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFPPFRNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFPPFRNQVILTAPNRS 418
QY 419 SYRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQDICYAPLND 478
Db 419 SYRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQDICYAPLND 478
QY 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
QY 539 ALSQADYCAPVPPFLAIGYKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAFLE 598
Db 539 ALSQADYCAPVPPFLAIGYKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAFLE 598
QY 599 EMRAFQRMAGMPQVTFPMAERSLEDEINRTTARDLPIFATSYIVIFLYISLALGYSYSS 658
Db 599 EMRAFQRMAGMPQVTFPMAERSLEDEINRTTARDLPIFATSYIVIFLYISLALGYSYSS 658
QY 659 RVWVDSKATLGGGAVVLGAVMAAGFPFSGYLSIRSVILQVVPVLVSVGADNIFIV 718
Db 659 RVWVDSKATLGGGAVVLGAVMAAGFPFSGYLSIRSVILQVVPVLVSVGADNIFIV 718
QY 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGCALTPMPAVRTFALTSL 778
Db 719 LEYQRLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGCALTPMPAVRTFALTSL 778
QY 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVKPQBLPPGQEGILLGFQKAYAP 838
Db 779 AVILDFLLQMSAFVALLSDSKRQEARLDVCCVKPQBLPPGQEGILLGFQKAYAP 838
QY 839 FLLHWITRGVV---LFLALFGVSLYSMCHISVGLDQELALPKDSYLLDLPFLNRYF 894
Db 839 FLLHWITRGVVPSQLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDLPFLNRYF 898
QY 895 EVGAPVYFTTLYGYNFSSAGMNAICSSAGCNFSPFTQKIQVATEPPEQSYLAIPASSSV 954
Db 895 EVGAPVYFTTLYGYNFSSAGMNAICSSAGCNFSPFTQKIQVATEPPEQSYLAIPASSSV 958
QY 955 DDFIDWLTSPSSCRLYISGNPKDPCFTVNSLNCNKCWSITMGSVRPSVBFHXYLPW 1014
Db 955 DDFIDWLTSPSSCRLYISGNPKDPCFTVNSLNCNKCWSITMGSVRPSVBFHXYLPW 1018
QY 1015 FLNDRNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAA 1074
Db 1015 FLNDRNIKCPKGLAAYSTSVNLTSQGVLASRFMAYHKPKNSQDYTEALRAARELAA 1078
QY 1075 NITADLRKVPDTPAPEVEPYTITNVFYEOYLITLPEGLFMLSCLVPTFAVSCILLGD 1134
Db 1075 NITADLRKVPDTPAPEVEPYTITNVFYEOYLITLPEGLFMLSCLVPTFAVSCILLGD 1138
QY 1135 LRSGLLNLLSIYMLIVDTVGFMAWDIISYNAVSLNLSAVGMSVEFVSHITRSPASTK 1194
Db 1135 LRSGLLNLLSIYMLIVDTVGFMAWDIISYNAVSLNLSAVGMSVEFVSHITRSPASTK 1198
QY 1195 PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLTLLGLLHG 1254
Db 1195 PTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLTLLGLLHG 1258
QY 1255 LVFLPVILSYVGPVDNVPALALQKRAEEA-----VAAMVASCNPSPSVST 1301
Db 1255 LVFLPVILSYVGPVDNVPALALQKRAEEGSGSHGGLAQITPPSPQLTTSNST 1313
```

RESULT 33

```
US-10-450-763-53052
; Sequence 53052, Application US/10450763
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 53052
; LENGTH: 1344
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (901)...(917)
; OTHER INFORMATION: OESTROGEN RECEPTOR SIGNATURE domain identified by eMATRIX,
; OTHER INFORMATION: accession number PR00543D, p-value=1.355e-09, raw score of 10.87
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1344)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-450-763-53052
```

```
Query Match 94.6%; Score 6536; DB 34; Length 1344;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1271; Conservative 7; Mismatches 21; Indels 16; Gaps 4;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN 60
Db 1 MAEAGLRGWLWALLRLAQSEPYTTHQPGYCAFYDECGKNPGLSGSLMTLSNVCSLSN 60
QY 61 TPARKITGDHLLLOKICPRLYTGNPTQACCSAKQLVSLBASISITKALLTRCPACSDNF 120
Db 61 TPARKITGDHLLLOKICPRLYTGNPTQACCSAKQLVSLBASISITKALLTRCPACSDNF 120
QY 121 VNLHCHNTCSPNQSLFINVTRVAQIGAGOLPAVVAEAEVQHSFAEQSYDSCSRVPA 180
Db 121 VNLHCHNTCSPNQSLFINVTRVAQIGAGOLPAVVAEAEVQHSFAEQSYDSCSRVPA 180
QY 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
Db 181 ATLAVGTWCGVYGSALCNAQRWLNFGDGTGNGLAPLDITFHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNSQDDVATCSQDCAAACP--AIARPOALDSTFYLGQMPGSLVLIILCSVFVV 298
Db 241 ARCNSQDDVATCSQDCAAACPXPAPRPSTPSTW--RQMPGSLVLIILCSVFVV 298
QY 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVL 358
Db 299 TILLVGFVRVAPARDKSMVDPKKGTSLSDKLSFSTHTLLGQFFQGGTGWASWPLTILVL 358
QY 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFPPFRNQVILTAPNRS 418
Db 359 SVIPVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHFPPFRNQVILTAPNRS 418
QY 419 SYRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQDICYAPLND 478
Db 419 SYRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNISLQDICYAPLND 478
QY 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
Db 479 NTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDG TAL 538
QY 539 ALSQADYCAPVPPFLAIGYKGYSEAEALIMTFLSNYPAGDPRLAQAKLWEAFLE 598
```



```
Db 539 ALSCHADYGPVFPFLALGGYKDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLE 598
Qy 599 ENRAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWS 658
Db 599 ENRAFQRMAGMFQVFTFAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWS 658
Qy 659 RVWDSKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSVLIOVVPFLVLSVGAADNIFV 718
Db 659 RVWDSKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSVLIOVVPFLVLSVGAADNIFV 718
Qy 719 LEYQRLPRRPGPREVHIGRALGRVAPSWMLLCSLSEALCFFLGALTMPAVRTFALTSG 778
Db 719 LEYQRLPRRPGPREVHIGRALGRVAPSWMLLCSLSEALCFFLGALTMPAVRTFALTSG 778
Qy 779 AVILDFLLQMSAFVALLSLSKQASRLDVCCCKVQOELPPPGQEGGLLGGFFOKAYAP 838
Db 779 AVILDFLLQMSAFVALLSLSKQASRLDVCCCKVQOELPPPGQEGGLLGGFFOKAYAP 838
Qy 839 FLHHWITRGVV-----LLFLALFGVLSYMSCHISVGLQOELALPKDSYLLDYFLNRYF 894
Db 839 FLHHWITRGVVPSQLLFLALFGVLSYMSCHISVGLQOELALPKDSYLLDYFLNRYF 898
Qy 895 EVGAPVYFTVLGYNFSSEAGMNAICSSAGCNPFSFTOKIYATFPPQSYLAIPASSWV 954
Db 899 EVGAPVYFTVLGYNFSSEAGMNAICSSAGCNPFSFTOKIYATFPPQSYLSLPASSWV 958
Qy 955 DDFDMLTPSSCCRLYISGPNKDFCPTSVNSLNCNKCMSITMGSVRPSVEQPHKYL 1014
Db 959 DDFDMLTPSSCCRLYISGPNKDFCPTSVNSLNCNKCMSITMGSVRPSVEQPHKYL 1018
Qy 1015 FLNDRPNIKCPKGLAASTSVNLSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAA 1074
Db 1019 FLNDRPNIKCPKGLAASTSVNLSDGQVLASRFMAYHKPLKNSQDYTEALRAARELAA 1078
Qy 1075 NITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLD 1134
Db 1079 NITADLRKVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLD 1138
Qy 1135 LRSGLLNLSIYMLVDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITRSPAISTK 1194
Db 1139 LRSGLLNLSIYMLVDTVGFMAWGISYNAVSLINLVSAGVMSVEFVSHITRSPAISTK 1198
Qy 1195 PTWLERAKETISMGSAVAGVAMTNLPGLIIVLGLAKAQLIOIPFRLMLLITLGLLHG 1254
Db 1199 PTWLERAKETISMGSAVAGVAMTNLPGLIIVLGLAKAQLIOIPFRLMLLITLGLLHG 1258
Qy 1255 LVFLPVILSYGVPDVPNALLAQKRAEEA-----VAAVMVASCPNHPRSVST 1301
Db 1259 LVFLPVILSYGVPDVPNALLAQKRAEEGGSGSHGGLAQIITPPSPQLTTSKST 1313
```

RESULT 34

```
PCT-US03-22467-2
; Sequence 2, Application PC/TUS0322467
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JBO1603-K-WI
; CURRENT APPLICATION NUMBER: PCT/US03/22467
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
PCT-US03-22467-2
```

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

```
Qy 1 MAEAGRLWLLALLRLAQSEPYTTIHOPGYCAFYDECGKNPBLSGSLMTLSNVCSLSN 60
Db 1 MAEAWL-GWLLWALLLSAQGELYTPRHEAGVCTFYEECGKNPBLSGSLTSLSNVCSLSN 59
Qy 61 TPARKITGDHLILLOKICPRLYTCGNTQ-ACCSAKQLVSLBASISITKALLTRCPACSDN 119
Db 60 TPARHVTGEHLALLQRICPRLYNGPNTTFFACSTKQLLSLESSMSITKALLTRCPACSDN 119
Qy 120 FVNLCHNTCPNQSLFINTRVAQLGAGOLPAVAYEAFYQHSFAEOSYDSCSVRVPA 179
Db 120 FVSLCHNTCPDQSLFINTRVVERGAGEPPAVVAYEAFYQRSFAEKAYESCQVRIPA 179
Qy 180 AATLAVGTMCGVYSALCNAQRLNFOGDTGNGLAPLDITFHLLEPPGQAVSGSIGPLNEG 239
Db 180 AASLAVGSMCGVYSALCNAQRLNFOGDTGNGLAPLDITFHLLEPGQALPDGLOPLNGK 239
Qy 240 VARNESQGDVATCSCQDCAASCPAIARPOALDSTFYLQMPGSLVLIILCSVAVVT 299
Db 240 IAPCNESQGDSDAVCSQDCAASCPIVPPPEALRPSFYMGMRPGWMLIIITFAVFLLS 299
Qy 300 ILLVGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFQGWGTWVASHPLTILVLS 359
Db 300 AVLVLRYVSNRNKKAEGPOEAPKLPHKHLSPHTILGRFQNMGTVAWSPLTILVLS 359
Qy 360 VIPVVALAAGLVFTELTTDPVELWSAPNSQARSEKAPHDQHPFPFRTNOVILTAPNRSS 419
Db 360 FIVVVALAAGUTFIETTTDPVELWSAPKQARKEKSFDEHFGPFPFRTNQVFTARNSS 419
Qy 420 YRYDSLLGPNKFGSGLDLDLLELLELQERLRLHQLQVMSPEAQRNISLQDICYAPLNPN 479
Db 420 YKYDSLLGSKNFGSGLSLDLELLELLELQERLRLHQLQVMSPEARNISLQDICYAPLNPN 479
Qy 480 TSLDCCINSLLQFPQNNRTLLLTANQTLNGQTSQVDWKQHFLYCANAPLTFKQGTALA 539
Db 480 TSLDCCVNSLLQFPQNNRTLLMLTANQTLNGQTSQVDWKQHFLYCANAPLTFKQGTALA 539
Qy 540 LSCMADYCAPVPFLAIGYKGYSEABALIMFSLNNYPAGDPRLAQAKLWEEAFLE 599
Db 540 LSCMADYCAPVPFLAVGQYQTDYSEABALIIIFSLNNYPADDPRMAQAKLWEEAFLE 599
Qy 600 MRAFORMAGMFQVFTFAERSLEDEINRTTAEDLPFATSYVIFLYISLALGSYSWSR 659
Db 600 MESFORNTSDKFQVAFSAERSLEDEINRTTQDLVPFAVSYIIVFLYISLALGSYSRCSR 659
Qy 660 VMWDSKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSVLIOVVPFLVLSVGAADNIFV 719
Db 660 VAVESKATLGLGGVAVVLGAVMAAGPFYSYLGIRSSVLIOVVPFLVLSVGAADNIFV 719
Qy 720 EYQRLPRRPGPREVHIGRALGRVAPSWMLLCSLSEALCFFLGALTMPAVRTFALTSG 779
Db 720 EYQRLPRRPGPREVHIGRALGRVAPSWMLLCSLSEALCFFLGALTMPAVRTFALTSG 779
Qy 780 VILDFLLQMSAFVALLSLSKQASRLDVCCCKVQOELPPPGQEGGLLGGFFOKAYAP 839
Db 780 IILDFLLQMTAFVALLSLSKQASRPDLVCCPSTRKLPPEKKEGILLRFRKIYADF 839
Qy 840 LLHWITRGVWLLFLALFGVLSYMSCHISVGLQOELALPKDSYLLDYFLNRYFVGP 899
Db 840 LLHREIRPVWMLLFTLFANLYLNCINVGDLQOELALPKDSYLLDYFLNRYFVGP 899
Qy 900 VYFVTTLGYNFSSEAGMNAICSSAGCNPFSFTOKIYATFPPQSYLAIPASSWVDDFD 959
Db 900 VYFVTTSYGNFSSEAGMNAICSSAGCKSFLTKIYASEFFDQSYVAIAGSWVDDFD 959
Qy 960 WLTP-SSCCRLYISGPNKDFCPTSVNSLNCNKCMSITMGSVRPSVEQPHKYL 1018
Db 960 WLTPSSCCRLYIRGPHKDEFCPTDTSFNCNKCNMRTLGPVRPTAEQPHKYL 1019
Qy 1019 RNVIKCPKGLAASTSVNLSDGQVLASRFMAYHKPLKNSQDYTEALRAARELANITA 1078
Db 1020 PNNIRCPKGLAASTSVNLSDGQVIAQFMAHYHKPLKNSQDYTEALRAARELANITA 1079
Qy 1079 DURKVPGTDPAPFVPPYITINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRS 1138
```


APPLICANT: Garcia-Calvo, Margarita
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF IDENTIFYING LIGANDS THEREOF
; FILE REFERENCE: 36134-PCT 074669.0111
; CURRENT APPLICATION NUMBER: PCT/US05/01469
; PRIOR FILING DATE: 2005-01-25
; PRIOR APPLICATION NUMBER: US 60/537,341
; PRIOR FILING DATE: 2004-01-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
PCT-US05-01469-2

Query Match 78.5%; Score 5421.5; DB 1; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKNPELSGSLMTLSNVSCLSN 59
61 TPARKITGDHLILLOKICPRLYTGPNTO-ACCSAKQLVLEASLSITKALLTRCPACSDN 119
DB 60 TPARKITGDHLILLOKICPRLYTGPNTO-ACCSAKQLVLEASLSITKALLTRCPACSDN 119
QY 120 FVNLCHNTCSNQSFLNTRVAVQAGQLPAAVAYAEFYQHSFABQSDYSCSRVUPA 179
DB 120 FVNLCHNTCSNQSFLNTRVAVQAGQLPAAVAYAEFYQHSFABQSDYSCSRVUPA 179
QY 180 AATLAVGTGCVYGSALCNAQRLNFGQDTGNGLAPLDTIFHLHPGQAGVSGIOPLNKG 239
DB 180 AATLAVGTGCVYGSALCNAQRLNFGQDTGNGLAPLDTIFHLHPGQAGVSGIOPLNKG 239
QY 240 VARNESQDDVATCSQDCAASCPAIPALPQALDSTFYLGQMPGSLVLIILCSYFVVT 299
DB 240 IAPCNESQDDSAVCSQDCAASCPVIPPPEALRPSFYMGPMGLALIIITFAVFVLLS 299
QY 300 ILLVGFVAPADKSNWDPKGTSLSKLSSTHTLLGQFQCGWGTWASHPLIIVLS 359
DB 300 AVLRLRVVSNRKNKAEGQEPAPKLPHKHLSPTHTILGRFFQNMGTTRVASHPLIIVLS 359
QY 360 VLPVVALAAGLFTLTTPDVELMSAPNSQARSEKAFDOHFGPPERTNOVLTAPNRSS 419
DB 360 FIVVTLAAGLFTLTTPDVELMSAPNSQARSEKAFDOHFGPPERTNOVLTAPNRSS 419
QY 420 YRYSLLGPKNFSGILDLDLLELELQERLRLHQLVNSPEAQRNISLODICYAPLNPDN 479
DB 420 KYDLSLLGSKNFSGILSDFLLELELQERLRLHQLVNSPEAQRNISLODICYAPLNPN 479
QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKQGTALA 539
DB 480 TSLSDCCVNSLLQYFQNNRTLLLTANQTLMGQTSQVDKDHFLYCANAPLTFKQGTSLA 539
QY 540 LSCMDYGPAPVPFPLAIGVKGQYSEARALMTFSLNNYPAGDPLRAQAKLWEAFLE 599
DB 540 LSCMDYGPAPVPFPLAIGVKGQYSEARALMTFSLNNYPAGDPLRAQAKLWEAFLE 599
QY 600 MRAFORMAGMQVTFPABRSLEDEINRTTAEPLPFIATSYIVIFYLSLALGSYSNR 659
DB 600 MESFORNTSKQVAFSAERSLEDEINRTTIQDLPVFVSVIYVIFYLSLALGSYSRCSR 659
QY 660 VMVDSKATILGQVAVLVAVMAAMGPFYSLGRSSILVILQVPPFLVLSVGADNPIFVYL 719
DB 660 VAVESKATILGQVAVLVAVMAAMGPFYSLGVPSLVIQVPPFLVLAAGADNPIFVYL 719
QY 720 EYQRLPRRCEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 779
DB 720 EYQRLPRRCEPREVHIGRALGRVAPSMLLCSLSEACFPFLGALTMPAVRTPALTSGLA 779
QY 780 VILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPFGQSGELLGPFQKAYAPP 839
DB 780 VILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPFGQSGELLGPFQKAYAPP 839

DB 780 IILDFLLQMSAFVALLSLDSKRQASRLDVCCVKPQELPPFGQSGELLGPFQKAYAPP 839
QY 840 LLHWITRGVLLFLALFGVSLYSNCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAP 899
DB 840 LLHRFIRVVMVLLFLTLFGANLYMCNINVGLDQELALPKDSYLLDYFLFLNRYFEVGAP 899
QY 900 VYFVTTGLGYNFSSAGMNAICSSAGCNMFSTFKIQYATEPPEQSYLAIPASSWVDDFD 959
DB 900 VYFVTTGYNFSSAGMNAICSSAGCNMFSTFKIQYATEPPEQSYLAIPASSWVDDFD 959
QY 960 WLTP-SSCCRLYISGPNKDFCPSVTNSLNCILKNCWSTMGSVRPSVQFHKYLPWFND 1018
DB 960 WLTPSSCCRLYIRGPHKDFCPSVTNSLNCILKNCWSTMGSVRPSVQFHKYLPWFND 1019
QY 1019 RPNIKCPKGGAAVSTSVNLTSDGOVLASRFMAVHKLKNSQDYTEALRAARELANITA 1078
DB 1020 PPNRCPKGGAAVSTSVNLTSDGOVLASRFMAVHKLKNSQDYTEALRAARELANITA 1079
QY 1079 DLKVPKGTDPAPFVPPYTTINVFYEQYLTILPEGLFMLSCLIVPTFAVSLGLDLRSG 1138
DB 1080 DLKVPKGTDPAPFVPPYTTINVFYEQYLTILPEGLFMLSCLIVPTFAVSLGLDLRSG 1139
QY 1139 LNLLSIVMILVDTVGFMAWDISYNVSLINLVSAGVMSVEFVSHITRSPAIKPTWL 1198
DB 1140 LNLLSIVMILVDTVGFMAWDISYNVSLINLVSAGVMSVEFVSHITRSPAIKPTWL 1199
QY 1199 ERAKEATISMSGSAVAGVAMTNLPGILVLGLAKAQLIQLFFPRMLLITLLGLHGLVFL 1258
DB 1200 ERAKDATVFMGSAVAGVAMTNLPGILVLGLAKAQLIQLFFPRMLLITLLGLHGLVFL 1259
QY 1259 PVILSYVGPDPNPALEQKRAEAAVAAVMVASCPNHFSPVSTADNIYVNHSEFSGI-KG 1317
DB 1260 PVILSYVGPDPNPALEQKRAEAAVAAVMVASCPNHFSPVSTADNIYVNHSEFSGI-KG 1316
QY 1318 AGATSNFLPNNGRQF 1332
DB 1317 ANAARSLPKSDQKF 1331

RESULT 37

US-10-621-758A-2
; Sequence 2, Application US/10621758A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K-US
; CURRENT APPLICATION NUMBER: US/10/621,758A
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-621-758A-2

Query Match 78.5%; Score 5421.5; DB 36; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

QY 1 MAEAGLRGWLWALLRLAQSEPYTTIHQPGYCAFDECGKNPELSGSLMTLSNVSCLSN 60
DB 1 MAAAWL-GWLLWALLLSAAQGLYTPKHEAGVCTFYEECGKNPELSGSLMTLSNVSCLSN 59
QY 61 TPARKITGDHLILLOKICPRLYTGPNTO-ACCSAKQLVLEASLSITKALLTRCPACSDN 119
DB 60 TPARKITGDHLILLOKICPRLYTGPNTO-ACCSAKQLVLEASLSITKALLTRCPACSDN 119

QY 120 FVNLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
Db 120 FVSLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
QY 180 AATLAVGTWCGYVGSALCNAQWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQLNKG 239
Db 180 AASLAVGSMCGYVGSALCNAQWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQLNKG 239
QY 240 VARCNESQGDVATCSQDCAASCPAIAIPQALDSTFYLGQMPGSLVLIILCSVPVAVT 299
Db 240 IAPCNESQGDVATCSQDCAASCPVIPPPEALRPSFYNGRMPGSLVLIILCSVPVAVT 299
QY 300 ILLVFRVAPARDKSMVDPKKGTSLSDKLSSTHTLLGQPCQGWGTWASWPLTILVLS 359
Db 300 AVLVRLRVVNRNKNKAEGPQEAAPKLPHKHLSPTHTILGRPFQNGWGTWASWPLTILVLS 359
QY 360 VTPVVALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHDPGPPFRNTQVILTPANRNS 419
Db 360 FIVVIALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHDPGPPFRNTQVILTPANRNS 419
QY 420 YRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNLSQDICYAPLNPN 479
Db 420 YKYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNLSQDICYAPLNPN 479
QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTFKDGTLA 539
Db 480 TSLSDCCVNSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTFKDGTLA 539
QY 540 LSCMADYGAPVPFELAIKGYKGVSEAEALITWTSLNYPAGDPLAOKLWEAFLE 599
Db 540 LSCMADYGAPVPFELAIKGYKGVSEAEALITWTSLNYPAGDPLAOKLWEAFLE 599
QY 600 MEAFORRMAGMFOVFTTAEBSLEDEINRTTARDLPFATSYIVIFLYISLALGSYSWSR 659
Db 600 MESFORRTSDKQVAFSAERSLEDEINRTTARDLPFATSYIVIFLYISLALGSYSWSR 659
QY 660 VMVDSKATLGLGGVAVLVGAVMAAGFFSYLGRSLVTLQVVPFVLVSGVADNIPFVL 719
Db 660 VAVESKATLGLGGVAVLVGAVMAAGFFSYLGRSLVTLQVVPFVLVSGVADNIPFVL 719
QY 720 EYQRLPRRGEPRVHIGRALGRVAPSMLLCSISEAICFPLGALTPMPAVRTPALTSGLA 779
Db 720 EYQRLPRRGEPRVHIGRALGRVAPSMLLCSISEAICFPLGALTPMPAVRTPALTSGLA 779
QY 780 VILDFLLQMSAFVALLSLDSKQEAERLSDVCCVPPQELPPQEGGLLGFQKAYAPF 839
Db 780 IILDFLLQMTAFVALLSLDSKQEAERLSDVCCVPPQELPPQEGGLLGFQKAYAPF 839
QY 840 LLHWITRGVLLLFALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLNRYFEVGP 899
Db 840 LLHRRFIRPVVMLLFLTPGALVLMCNINVLQDQELALPKDSYLLDYFLNRYFEVGP 899
QY 900 VYFVTLGNFSSAGMNAICSSAGCNWFSFKQYQATEPPEQSYLAI PASWDDDFD 959
Db 900 VYFVTLGNFSSAGMNAICSSAGCNWFSFKQYQATEPPEQSYLAI PASWDDDFD 959
QY 960 WUTP--SSCCRLVYSGNPKDPCSTVNSLNCNKNMSITWGSVRPSVEGFHYLWPLFND 1018
Db 960 WUTPSSCCRLVYSGNPKDPCSTVNSLNCNKNMSITWGSVRPSVEGFHYLWPLFND 1018
QY 1019 RPNIRCPKGLAAYTSVNLSDGQVLAERFMAHYHKLKNSODYTEALRAARELANITA 1078
Db 1020 RPNIRCPKGLAAYTSVNLSDGQVLAERFMAHYHKLKNSODYTEALRAARELANITA 1078
QY 1079 DLKVPDTPAPVEVPYITNVPYQYLTILPEGLFMLSCLVPTFPAVSCLLGLDLRSG 1138
Db 1080 DLKVPDTPAPVEVPYITNVPYQYLTILPEGLFMLSCLVPTFPAVSCLLGLDLRSG 1138
QY 1139 LNLNLSIVMLVDTVGFEMALWDISYNAVSLINLVAAGVMSVFVSHITSPAIKPTWL 1198
Db 1140 LNLNLSIVMLVDTVGFEMALWDISYNAVSLINLVAAGVMSVFVSHITSPAIKPTWL 1198
QY 1199 ERAKEATISMGSAVFAGVAMTNPGLVILGLAKAQIQTFFRLNLLITLLGLLHGLVFL 1258

Db 1200 ERAKQATVPMGSAVFAGVAMTNPGLVILGLFAQAQIQTFFRLNLLITLLGLLHGLVFL 1259
QY 1259 PVILSVYGVDPNPALALEOKRAEEAAVAAVMVASCNHPHSRVSTADNIYVNHSPGSI-KG 1317
Db 1260 FVVLVYLGPDVNOALVOEKLASEA-AVAPEFSCTQYPSPADADAN--VNYGAPFELANG 1316
QY 1318 AGAINFLPNNGRQF 1332
Db 1317 ANAARSSLPKSDQK 1331
RESULT 38
US-10-646-301A-2
; Sequence 2, Application US/10646301A
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; FILE REFERENCE: JBO1603-KI-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-646-301A-2
Query Match 78.5%; Score 5421.5; DB 36; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
QY 1 MRAEGLRGWLLWALLRLAQSEPTIITIQPGYCAPVDECKNPELSGLMTLSNVSCLSN 60
Db 1 MAAAWL-GWLLWALLSAAQSGELYTPKHEAGVCTFYECGKNPELSGGITLSNVSCLSN 59
QY 61 TPARKITGDHLLILQKICPLRYTGENTO-ACCSAQQLVSLSEASLSTTKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPLRYTGENTO-ACCSAQQLVSLSEASLSTTKALLTRCPACSDN 119
QY 120 FVNLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
Db 120 FVSLHCHNTCSNQSIFINVTVAQIGAGOLPAVVAYEAFYQHSFAEQSYDSCSRVRPA 179
QY 180 AATLAVGTWCGYVGSALCNAQWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQLNKG 239
Db 180 AASLAVGSMCGYVGSALCNAQWLNFGQDTGNGLAPLDTITFHLLEPGQAVGSGIQLNKG 239
QY 240 VARCNESQGDVATCSQDCAASCPAIAIPQALDSTFYLGQMPGSLVLIILCSVPVAVT 299
Db 240 IAPCNESQGDVATCSQDCAASCPVIPPPEALRPSFYNGRMPGSLVLIILCSVPVAVT 299
QY 300 ILLVFRVAPARDKSMVDPKKGTSLSDKLSSTHTLLGQPCQGWGTWASWPLTILVLS 359
Db 300 AVLVRLRVVNRNKNKAEGPQEAAPKLPHKHLSPTHTILGRPFQNGWGTWASWPLTILVLS 359
QY 360 VTPVVALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHDPGPPFRNTQVILTPANRNS 419
Db 360 FIVVIALAAGLVFTLTDTPVELWSAPNSQARSEKAFHQHDPGPPFRNTQVILTPANRNS 419
QY 420 YRYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNLSQDICYAPLNPN 479
Db 420 YKYDSLLGPKNFSGILDLDLLELLELQERLHLQVMSPEAQRNLSQDICYAPLNPN 479
QY 480 TSLYDCCINSLLQYFQNNRTLLLTANQTLMGOTSDQVMDKHFLYCANAPLTFKDGTLA 539

```
Db 480 TSLDCCVNSLLQYFQNNRTLLMLTANQTLNGQTSLVMDKDHFLYCANAPLTFKDGTSLA 539
Qy 540 LSCMADYGAPVFPFLAIGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEAFLEE 599
Db 540 LSCMADYGAPVFPFLAIGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEAFLEE 599
Qy 600 MRAFORMMAGFQVFTFAERSLEDEINRTTADLPIPATSYIVIFLYISLALGSYSWSR 659
Db 600 MESFORNTSDKPFQVAFSAERSLEDEINRTTADLPIPATSYIVIFLYISLALGSYSRCSR 659
Qy 660 VMVDSKATLGLGGVAVLVGAVMAAGFFSGLIRSSVLQVVPFLVLSVGDNIFIFVL 719
Db 660 VAVESKATLGLGGVAVLVGAVLAAAGFYSLGVPSLVIQVVPFLVAVGADNIFIFVL 719
Qy 720 EYQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Db 720 EYQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Qy 780 VILDPELLQMSAFVALLSDSKQREASRLDVCCCVKQPELPPGQEGILLGFQKAYAPF 839
Db 780 VILDPELLQMTAFVALLSDSKQREASRPDLVCCFSTRKLPDPPEKEGGLLRFKRIYAPF 839
Qy 840 LIHWTIRGVLILFLALFGVSLYSCHISVGLDQELALPKDSYLLDYFLFLNRYEFGAP 899
Db 840 LLHRFIRPVMLFLTFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
Qy 900 VYFVTTLGVNFSSESAGMAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFID 959
Db 900 VYFVTTLGVNFSSESAGMAICSSAGCNPFSTQKIQYATEPPEQSYLAIPASSWDDFID 959
Qy 960 WLTP-SSCCRLYISGPNKDKFCPSTVNSLNCNKNMSITMGSVRPSVQFHKYLPWFLND 1018
Db 960 WLTPSSCCRLYIRGPHKDFECPSTDTSPNCLKNCNMTLGVPRTPASQFHKYLPWFLND 1019
Qy 1019 RNIKCPKGLAAYGTSVNLTDGQVLASRFMAYHKPLKNSODYTEALRAARELANITA 1078
Db 1020 PNIRKCPKGLAAYRTSVNLSDDGVIAQFVAYHKPLNSQDTEALRAARELANITA 1079
Qy 1079 DLKRVGTDPAFEVPPYITINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
Db 1080 DLKRVGTDNPEVPPYITINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDMCSG 1139
Qy 1139 LNLASIVMLVDTCFMAWDIISYNAVSLINLVSAGVSVFVSHITSPASTKPTWL 1198
Db 1140 ILNLSIIMLVDTTGLMAVWGISYNAVSLINLVTAVGMSVFEVSHITSPASTKPTWL 1199
Qy 1199 ERAKEATISMGSAVFAGVAMTNLPGILVLGAKAQLIQIFFFRLMLLITLLGLLHGLVPL 1258
Db 1200 ERAKATVPMGSAVFAGVAMTNLPGILVLGAKAQLIQIFFFRLMLLITLLGLLHGLVPL 1259
Qy 1259 PVLSYGVDPNPALEOKREAEVAAVMVASCPNHPRSRVSTADNIVNHSFEGSI-KG 1317
Db 1260 PVVLSYLGDPVNOALVOEKLASEA-AVAPESPQYPSPADADAN--VNYGFAPELAHG 1316
Qy 1318 AGAISNPLNNGROP 1332
Db 1317 ANAARSSLPKSDQKF 1331
```

RESULT 39

```
US-10-663-208A-2
; Sequence 2, Application US/10663208A
; GENERAL INFORMATION:
; APPLICANT: Altman, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603K2 US
; CURRENT APPLICATION NUMBER: US/10/663,208A
; CURRENT FILING DATE: 2003-09-16
; PRIOR APPLICATION NUMBER: 60/397,442
```

```
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-10-663-208A-2
```

Query Match 78.5%; Score 5421.5; DB 36; Length 1331;

Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;

```
Qy 1 MAEAGRLGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELGSGLMTLSNVSCLSN 60
Db 1 MAEAGRLGWLWALLRLAQSEPYTTIHQGYCAFYDECGKNPELGSGLMTLSNVSCLSN 59
Qy 61 TPARKITGDHLLILQKICPRLYTGPNTO-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 60 TPARKITGDHLLILQKICPRLYTGPNTO-ACCSAKQLVSLASLSITKALLTRCPACSDN 119
Qy 120 FVNLHCHNTCSFNOSLFINVTRVAQAGOLPVAVYEAFFYQHSFAEQSYDSCSRVRPA 179
Db 120 FVNLHCHNTCSFNOSLFINVTRVAQAGOLPVAVYEAFFYQHSFAEQSYDSCSRVRPA 179
Qy 180 AATLAVGTMCGVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVSGGIQPLNEG 239
Db 180 AATLAVGTMCGVGSALCNAQRWLNFGQDTGNGLAPLDITFHLLEPQAVSGGIQPLNEG 239
Qy 240 VARNESGDDVATVCSQDCAASCAPAIARPOALDSTFVLGOMPSGLVILILCSYPAVVT 299
Db 240 VARNESGDDVATVCSQDCAASCAPAIARPOALDSTFVLGOMPSGLVILILCSYPAVVT 299
Qy 300 ILLVGFVAPARDKSKMVDPKKTSLSDKLSFTHTLLGQFQGMGTWVASWPLTILVLS 359
Db 300 ILLVGFVAPARDKSKMVDPKKTSLSDKLSFTHTLLGQFQGMGTWVASWPLTILVLS 359
Qy 360 VIPVVALAAGLVFTLTTPVELMSAPNSQARSEKAFHDQHFGPPFRFNQVILTPARNSS 419
Db 360 VIPVVALAAGLVFTLTTPVELMSAPNSQARSEKAFHDQHFGPPFRFNQVILTPARNSS 419
Qy 420 YRYSLLILGPKNFSGILDLLELLELEQERLHLQVMSPEAQRNLSQDICYAPLNPDN 479
Db 420 YRYSLLILGPKNFSGILDLLELLELEQERLHLQVMSPEAQRNLSQDICYAPLNPDN 479
Qy 480 TSLYDCCINSLLQYFQNNRTLLMLTANQTLNGQTSLVMDKDHFLYCANAPLTFKDGTSLA 539
Db 480 TSLYDCCINSLLQYFQNNRTLLMLTANQTLNGQTSLVMDKDHFLYCANAPLTFKDGTSLA 539
Qy 540 LSCMADYGAPVFPFLAIGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEAFLEE 599
Db 540 LSCMADYGAPVFPFLAIGYKGKDYSEABALIMTFSLNNYPAGDPRLAQAKLWEAFLEE 599
Qy 600 MRAFORMMAGFQVFTFAERSLEDEINRTTADLPIPATSYIVIFLYISLALGSYSWSR 659
Db 600 MRAFORMMAGFQVFTFAERSLEDEINRTTADLPIPATSYIVIFLYISLALGSYSWSR 659
Qy 660 VMVDSKATLGLGGVAVLVGAVMAAGFFSGLIRSSVLQVVPFLVLSVGDNIFIFVL 719
Db 660 VMVDSKATLGLGGVAVLVGAVMAAGFFSGLIRSSVLQVVPFLVLSVGDNIFIFVL 719
Qy 720 EYQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Db 720 EYQRLPRPGEPREVHIGRALGRVAPSMMLCSLSAICFFLGALTPMPAVRTFALTSGLA 779
Qy 780 VILDPELLQMSAFVALLSDSKQREASRLDVCCCVKQPELPPGQEGILLGFQKAYAPF 839
Db 780 VILDPELLQMTAFVALLSDSKQREASRPDLVCCFSTRKLPDPPEKEGGLLRFKRIYAPF 839
```

QY 840 LLHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAP 899
DB 840 LLHFRIRPVVMLLFLTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
QY 900 VYFVTTLVGNFSSAGNATCSSAGCNPFSTQIOYATEPFGQSYLAIPASSWVDDFID 959
DB 900 VYFVTTSGNFSSEAGNATCSSAGCKSFSLTQIOYASEFPDQSVVAIAASSWVDDFID 959
QY 960 WLTP-SGCCRLYISGPNKDKPCFSTVNSLNCNKMISITMGSVRPSVEQPHKYLWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCSTDTSPNCLKNCMRTLGPVPTAEQPHKYLWFLND 1019
QY 1019 RNKICPKGGLAAYSTSVNLTSQGLASRPMAYHKLKNSQDYTALRAARELANITA 1078
DB 1020 PPNIRCPKGLAAYSTSVNLSSDQGVIAQFMAYHKLKNSQDFTALRASRLAANITA 1079
QY 1079 DLKRVGTDPAFVFPYTTINVEYQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSG 1138
DB 1080 DLKRVGTDENFEVFPYTTINVEYQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSG 1139
QY 1139 LNLISVIMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198
DB 1140 LNLISVIMLVDTIGLMAWGISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1199
QY 1199 ERAKEATISGSAFVAGVAMTNLPGLILVLGLAKAQLIQIFFFRLNLLITLGLLHGLVFL 1258
DB 1200 ERAKDATVFMGSAFVAGVAMTNFPGLILVLGLAKAQLIQIFFFRLNLLITLGLLHGLVFL 1259
QY 1259 PVLTSVGVDPNPALALEKRAEAAVAVVAVSCPNHPSRVSTADNINVHSPESI-KG 1317
DB 1260 PVLTSVGVDPNPALVQESKLAESA-AVAPESPQYPSPADADAN--VNYGAPELAHG 1316
QY 1318 AGAISNFPNNGRQF 1332
DB 1317 ANAARSSLPKSDQKF 1331

RESULT 40

US-10-736-769-2
; Sequence 2, Application US/10736769
; GENERAL INFORMATION:
; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: J01603-K3-US
; CURRENT APPLICATION NUMBER: US/10/736,769
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1331
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-736-769-2

Query Match 78.5%; Score 5421.5; DB 37; Length 1331;
Best Local Similarity 77.8%; Pred. No. 0;
Matches 1039; Conservative 120; Mismatches 169; Indels 7; Gaps 6;
QY 1 MAEAGLRGWLWALLIRLAQSEPTTIHQPGYCAFYDECKRNPGLSGSLMTLSNVSCLSN 60
DB 1 MAAAML-GWLLWALLSAAQGLYTPKHBAGVCTFYBECGRNPGLSGSLMTLSNVSCLSN 59

QY 61 TPARKITGDHLLILLQKICPRLYTGPNQ-ACCSAKQLVLSLEASLSITKALLTRCPACSDN 119
DB 60 TPARKITGDHLLILLQKICPRLYTGPNQ-ACCSAKQLVLSLEASLSITKALLTRCPACSDN 119
QY 120 FVNLHCHNTCSNQSIFINVTVAQIAGOLPAVAYEAPYQHSFABEQSDVDSRVRVPA 179
DB 120 FVNLHCHNTCSNQSIFINVTVAQIAGOLPAVAYEAPYQHSFABEQSDVDSRVRVPA 179
QY 180 AATLAVTGMCGVYVSALCNAQRLNFGDGTGNGLAPLDTIFHLLLEPQAVGSGIQPLNEG 239
DB 180 AATLAVTGMCGVYVSALCNAQRLNFGDGTGNGLAPLDTIFHLLLEPQAVGSGIQPLNEG 239
QY 240 VARNESQDQDVATCSQDCAASCPAIPQALDSTFYLGOMPGSLVLIIILCSVFAVVT 299
DB 240 IAPCNESQDQDASVCSQDCAASCPVIPPPEALRPSFYMGMPGMLALIIITAVFVLLS 299
QY 300 ILLVGRFVAPADKSNVDPKKGTSLSDKLSSTHTLLGQFFQCGWGTWASWPLTILVLS 359
DB 300 AVLRVRVVSNNKNAEGQBPAPKLPKHKLSPHTILGRFFQNGWTRVASWPLTILVLS 359
QY 360 VIPVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTQVILTAPNRSS 419
DB 360 FIVVIALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTQVILTAPNRSS 419
QY 420 XYDLSLLGPKNPSGILDLDLLELELQERLHLQVWSPQARNISLQDICYAPLNPDN 479
DB 420 XYDLSLLGPKNPSGILDLDLLELELQERLHLQVWSPQARNISLQDICYAPLNPDN 479
QY 480 TSLYDCCINSILQYFQNNRTLLLTANQILMGTSQVNDKDHPLYCANAPLTFKDGTALA 539
DB 480 TSLYDCCINSILQYFQNNRTLLLTANQILMGTSQVNDKDHPLYCANAPLTFKDGTALA 539
QY 540 LSCMADYGAPVPFPLAIGYKGDYSEARALTWPSLNYPAGDPRLAQAQKWEAPLE 599
DB 540 LSCMADYGAPVPFPLAIGYKGDYSEARALTWPSLNYPAGDPRLAQAQKWEAPLE 599
QY 600 MEAFQRRMAGMFQVTFPERSLEDEINRTABDLPFATSYIYIFLYISALGSYSWSR 659
DB 600 MESFORNTSDKQVAFSAERSLEDEINRTIQDLPVFAVSYIYIFLYISALGSYSWSR 659
QY 660 VMDSKATLGLGVAVVILGAVMAAGFPFSLGIRSLVILQVVPFVLVSVGANIIFVL 719
DB 660 VAVESKATLGLGVAVVILGAVMAAGFPFSLGIRSLVILQVVPFVLVSVGANIIFVL 719
QY 720 EYQRLPRRGPBPREVHIGRALGRVAPSMILCSISRAICPFLGALTMPAVRTPALTSGLA 779
DB 720 EYQRLPRRGPBPREVHIGRALGRVAPSMILCSISRAICPFLGALTMPAVRTPALTSGLA 779
QY 780 VILDFLLQMSAFVALLSLDSKQESRLDVCCVQPELPPQGGEGLLIGFPQKAYAPF 839
DB 780 IILDFFLQMTAFVALLSLDSKQESRLDVCCVQPELPPQGGEGLLIGFPQKAYAPF 839
QY 840 LLHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFVGVAP 899
DB 840 LLHFRIRPVVMLLFLTLFGANLYLMCNINVLGDLQELALPKDSYLLDYFLFLNRYLEVGP 899
QY 900 VYFVTTLVGNFSSAGNATCSSAGCNPFSTQIOYATEPFGQSYLAIPASSWVDDFID 959
DB 900 VYFVTTSGNFSSEAGNATCSSAGCKSFSLTQIOYASEFPDQSVVAIAASSWVDDFID 959
QY 960 WLTP-SGCCRLYISGPNKDKPCFSTVNSLNCNKMISITMGSVRPSVEQPHKYLWFLND 1018
DB 960 WLTPSSCCRLYIRGPHKDEFCSTDTSPNCLKNCMRTLGPVPTAEQPHKYLWFLND 1019
QY 1019 RNKICPKGGLAAYSTSVNLTSQGLASRPMAYHKLKNSQDYTALRAARELANITA 1078
DB 1020 PPNIRCPKGLAAYSTSVNLSSDQGVIAQFMAYHKLKNSQDFTALRASRLAANITA 1079
QY 1079 DLKRVGTDPAFVFPYTTINVEYQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSG 1138
DB 1080 DLKRVGTDENFEVFPYTTINVEYQVLTILPGLFMLSICLVPTFAVSCLLGLDLRSG 1139
QY 1139 LNLISVIMLVDTVGFMAWDISYNAVSLINLVSAGMSVEFVSHITRSPAISTKPTWL 1198


```
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US03-22467-12

Query Match      78.3%; Score 5407; DB 1; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGHLLWALLRLLAQSEPTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLNTPA 63
DB 3 AAQOGLLWALLNSAQGLYPTTHKAGCTFYECGKNPELSGSLMTLSNVSCLNTPA 62

QY 64 RKITGDHLLILLOKICPRLYTGN-TOACCSAKQLVSLBASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGDHLLALQVCPRLYNGPNDTYACCSTQQLVSLDSSLITKALLTRCPACSNFVS 122

QY 123 LHCHNTCSNQSILFNVTRVAQAGQLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAT 182
DB 123 IHCHNTCSNQSILFNVTRVQDQQLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAS 182

QY 183 LAVGTGCVGYGALCNAQWLNFPQDGTGNGLAPLDTIFHLLEPQAVGSGIOPLENGVAR 242
DB 183 LAVGSMCGVYGALCNAQWLNFPQDGTGNGLAPLDTIFHLLEPQAVGSGIOPLENGVAR 242

QY 243 CNEQGDVVATCSQDCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSFPAVVITLL 302
DB 243 CNEQGEDSAAQSCQDCAASCPVIPPALRPSFTNGRMFGMLALIIITFAVFLSVVL 302

QY 303 VGRVPAPARDKSMVDPKKGTSLSKLSFTLLGQFGQMGWTWVWASPLTILVLSVIP 362
DB 303 VYLRVASNKNKTAGSQEAPNLPRKRFPSPHVLGRFPESNGTRVASWPLTVLALSFI 362

QY 363 VVALAAGLVTELTDPVLEWAPNSQARSEKAFHQDHPFPFRNQVILTPAPNSRYK 422
DB 363 VIALSVGLTFIELTDPVLEWAPNSQARSEKAFHQDHPFPFRNQVILTPAPNSRYK 422

QY 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482
DB 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482

QY 483 YDCINSLLOQYQNNRTLLLTANOTLMQTSQVQWQKHFLYCANAPLTFKDGTALALSC 542
DB 483 TDCCVNSLLQYQNNHTLLLTANQTLNGQTSILVDWKHFLYCANAPLTFKDGTALALSC 542

QY 543 MADYGAPVPPFLAIGYKGYKGYSEBALIMTFSNLNYPAGDPRLAQAKLWBEAFLEEMRA 602
DB 543 IADYGAPVPPFLAIGYKGYKGYSEBALIITFSINNPADDPMAHAKLWBEAFLEEMQS 602

QY 603 FORRMAGMFOVTFATERSLEDRNTTADIEDIFATSYIVIFLYTSLALGSSSSSRVWV 662
DB 603 FORSTADKFOIAFSAERSLEDRNTTIDQLPVFAISLYIVLYTSLALGSSSRVAV 662

QY 663 DSKATILGLGVAVLVGAVMAAGFFSYLGRSSVLTVQVVPFLVLSVGADNPIFVLEYQ 722
DB 663 DSKATILGLGVAVLVGAVMAAGFFSYLGVPSVLTVQVVPFLVLSVGADNPIFVLEYQ 722

QY 723 RLPRPGEPREVHIHGRALGRVAPSMMLCSLSBAICPFLGALTPMPAVRTFALTSGLAIVL 782
DB 723 RLPRPGEAREAHIGRTLGSVAPSMMLCSLSBAICPFLGALTPMPAVRTFALTSGLAIF 782

QY 783 DFLQMSAFVALLSLSKQESASRLDVCCVQDELPPPGQEGLLGLFFQKAYAPFLH 842
DB 783 DFLQMTAFVALLSLSKQESASRPDWCFCFSRNLPPPKQEGLLGLFFKRIYTPFLH 842

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYF 902
DB 843 RTRPVLLFLVLFGANLYLMCNISVGLDQDALPKDSYLLDYFLNRYLEVGPVYF 902

QY 903 VTTLGYNFSSEAGMNAICSSAGCNFNSFTQKIYATEFPPEQSYLAIPASSWVDDFDMLT 962
```

```
Db 903 DTTSGYNFSTEAGMNAICSSAGCBFSLTKIYASEFPNQSVATAASSWVDDFDMLT 962
QY 963 P-SSCCRLYISGPNKDKPCSTVNSLNCILKNCWSTIMGSRVPSVEQFHKYLPWFLNDRPN 1021
Db 963 PSSCCRIYTRGPHKDBFCFSTDTSFNCLKNCNRTLGVPVPTTEQFHKYLPWFLNDRPN 1022
QY 1022 IKCPKGGLAAYSTSVNLTSQDQVLASRFMAYHKPLKNSQDYTEALRAARELANITADLR 1081
Db 1023 IRCPKGGLAAYRTSVNLSDGQIIASQPMAYHKPLRNSQDTEALRASRLAANITABEL 1082
QY 1082 KVPGETDPAPFEPYITNVFYEOYLITLPEGLPMLSLCLVPTPAFVSCILGLDLRSLN 1141
Db 1083 KVPGETDPFEPYITNSFYQOYLTVLPEGFTTALCFVPTFWVCYLLGLDIRSGILN 1142
QY 1142 LLSIIVMILVDVTGFMALMDISYNVSLINLVSAGVMSVSVSHITRSFAISTKPTWLER 1201
Db 1143 LLSIIMILVDVTIGLMAVWGISYNVSLINLVTAVGMSVSVSHITRSFAISTKPTLER 1202
QY 1202 KBATISMGSAVPAGVAMTNLPGILVLGLAKAQLIQIFFRLNLLITLLGLHGLVLPV 1261
Db 1203 KDATIFMGSAVPAGVAMTNFPGILILGPAQAQLIQIFFRLNLLITLLGLHGLVLPV 1262
QY 1262 LSYVGDVNPALALQKRAEAAVAVMVASCPNHPSRVSTADNIYNHSEFES- IKGAGA 1320
Db 1263 LSYLGDVNOALVLEKLEATEA-AMVSEBPCQYPPADANTSDYVNYGFPETPEINA 1321
QY 1321 ISNLPNNGROF 1332
Db 1322 ASSSLPKSDQKP 1333

RESULT 43
PCT-US03-40113-12
; Sequence 12, Application PC/TUS0340113
; GENERAL INFORMATION:
; APPLICANT: Schering Corporation
; TITLE OF INVENTION: NPC1L1 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K3-WI
; CURRENT APPLICATION NUMBER: PCT/US03/40113
; CURRENT FILING DATE: 2003-12-16
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: 10/646,301
; PRIOR FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 10/663,208
; PRIOR FILING DATE: 2003-09-16
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
PCT-US03-40113-12

Query Match      78.3%; Score 5407; DB 1; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGHLLWALLRLLAQSEPTTIHQPGYCAFYDECGKNPELSGSLMTLSNVSCLNTPA 63
DB 3 AAQOGLLWALLNSAQGLYPTTHKAGCTFYECGKNPELSGSLMTLSNVSCLNTPA 62

QY 64 RKITGDHLLILLOKICPRLYTGN-TOACCSAKQLVSLBASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGDHLLALQVCPRLYNGPNDTYACCSTQQLVSLDSSLITKALLTRCPACSNFVS 122

QY 123 LHCHNTCSNQSILFNVTRVAQAGQLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAT 182
DB 123 IHCHNTCSNQSILFNVTRVQDQQLPAVVAEYAFYQHSFAEQSYDCSRRVPPAAAS 182

QY 183 LAVGTGCVGYGALCNAQWLNFPQDGTGNGLAPLDTIFHLLEPQAVGSGIOPLENGVAR 242
DB 183 LAVGSMCGVYGALCNAQWLNFPQDGTGNGLAPLDTIFHLLEPQAVGSGIOPLENGVAR 242

QY 243 CNEQGDVVATCSQDCAASCPAIAARPQALDSTFYLGQMPGSLVLIILCSFPAVVITLL 302
DB 243 CNEQGEDSAAQSCQDCAASCPVIPPALRPSFTNGRMFGMLALIIITFAVFLSVVL 302

QY 303 VGRVPAPARDKSMVDPKKGTSLSKLSFTLLGQFGQMGWTWVWASPLTILVLSVIP 362
DB 303 VYLRVASNKNKTAGSQEAPNLPRKRFPSPHVLGRFPESNGTRVASWPLTVLALSFI 362

QY 363 VVALAAGLVTELTDPVLEWAPNSQARSEKAFHQDHPFPFRNQVILTPAPNSRYK 422
DB 363 VIALSVGLTFIELTDPVLEWAPNSQARSEKAFHQDHPFPFRNQVILTPAPNSRYK 422

QY 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482
DB 423 DSIILGPKNFGSLDLDLLELLEQLERHLQVMSPEAQRNLSLODICYAPLNPDNTSL 482

QY 483 YDCINSLLOQYQNNRTLLLTANOTLMQTSQVQWQKHFLYCANAPLTFKDGTALALSC 542
DB 483 TDCCVNSLLQYQNNHTLLLTANQTLNGQTSILVDWKHFLYCANAPLTFKDGTALALSC 542

QY 543 MADYGAPVPPFLAIGYKGYKGYSEBALIMTFSNLNYPAGDPRLAQAKLWBEAFLEEMRA 602
DB 543 IADYGAPVPPFLAIGYKGYKGYSEBALIITFSINNPADDPMAHAKLWBEAFLEEMQS 602

QY 603 FORRMAGMFOVTFATERSLEDRNTTADIEDIFATSYIVIFLYTSLALGSSSSSRVWV 662
DB 603 FORSTADKFOIAFSAERSLEDRNTTIDQLPVFAISLYIVLYTSLALGSSSRVAV 662

QY 663 DSKATILGLGVAVLVGAVMAAGFFSYLGRSSVLTVQVVPFLVLSVGADNPIFVLEYQ 722
DB 663 DSKATILGLGVAVLVGAVMAAGFFSYLGVPSVLTVQVVPFLVLSVGADNPIFVLEYQ 722

QY 723 RLPRPGEPREVHIHGRALGRVAPSMMLCSLSBAICPFLGALTPMPAVRTFALTSGLAIVL 782
DB 723 RLPRPGEAREAHIGRTLGSVAPSMMLCSLSBAICPFLGALTPMPAVRTFALTSGLAIF 782

QY 783 DFLQMSAFVALLSLSKQESASRLDVCCVQDELPPPGQEGLLGLFFQKAYAPFLH 842
DB 783 DFLQMTAFVALLSLSKQESASRPDWCFCFSRNLPPPKQEGLLGLFFKRIYTPFLH 842

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDQELALPKDSYLLDYFLNRYFEVGPVYF 902
DB 843 RTRPVLLFLVLFGANLYLMCNISVGLDQDALPKDSYLLDYFLNRYLEVGPVYF 902

QY 903 VTTLGYNFSSEAGMNAICSSAGCNFNSFTQKIYATEFPPEQSYLAIPASSWVDDFDMLT 962
```


Db 603 FORSTADKQIARSAERSLEDEINRTTIQDLPVFAISYLIVFYISIALGSGSRWRSVAV 662
Qy 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGIRSSILVILQVVPFLVSVGADNIFIFVLEYQ 722
Db 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGIRSSILVILQVVPFLVSVGADNIFIFVLEYQ 722
Qy 723 RLPRRGEPREVHIGALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMPEGEAREHIGRTLGSVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Qy 783 DFLQMSAFVALLSLDSKQEARSLDVCCVCKPQELPPQCGEGLLGFKYAFKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQEARSDPVCCFSSRLNPPPKQKGLLCCFRKIYTFPLH 842
Qy 843 WITRGVLLFLFALFGVSYLSMCHISVGLDQELALPKDSYLDYFYLFLNRYFEVGPVVF 902
Db 843 RPIRPVLLFLVLFGANLYLMCNISVGLDQDLALPKDSYLDYFYLFLNRYFEVGPVVF 902
Qy 903 VTLGYNFSSAGMNAICSSAGCNFSPFTQKIYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIYASEFPNQSYVAIAASSWVDDFIDWLT 962
Qy 963 P-SSCCRLYISGPNKDKFCPTVNSLNCNKMSITWGSVRSPVQFHKYLPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFCEPSTDTSFNCLKNMRTLGPRVPTTEQFHKYLPWFLNDRN 1022
Qy 1022 IKCPKGLAAYSTVNLTSVGQVLAASRMAHYKPKNSQDYTEALRAARELANITADLR 1081
Db 1023 IKCPKGLAAYSTVNLSDGQIIASQFMAHYKPLNSQDFTEALRASRLANITAEUR 1082
Qy 1082 KVPGETDPAFVPPYITINVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLN 1141
Db 1083 KVEGTDNPEVPPYITISNVFYQYLTVPPEGITLALCPVTFVVCYLLGLDIRSGILN 1142
Qy 1142 LLSIVMLVDTGPMALNDISYNAVSLINLVSAVMSVEFVSHITRSPALSTKPTWLER 1201
Db 1143 LLSIIMILVDTGLMAVWGISYNAVSLINLVAVGMSVEFVSHITRSPAVSTKPTRLER 1202
Qy 1202 KEATISMSGSAFAGVANTNLPGLVLGLAKAOLIQIFFRMLNLLITLGLLHGLVPLVI 1261
Db 1203 KDATTFMGSAFAGVANTNFPGLILGLFAQAOLIQIFFRMLNLLITLGLLHGLVPLVI 1262
Qy 1262 LSYGVDVNPALALEOKRAEAAVAAVMVASCPNHPSSRVSTADNIYVNHSPFEGS-1KGAGA 1320
Db 1263 LSYLGDVNOALVLEBKATER-AWVSEPCQYPPADANTSDYVYGFNPEFPEINA 1321
Qy 1321 ISNPLNNGRQF 1332
Db 1322 ASSSLPKSQKF 1333

RESULT 45
PCT-US05-27579-3
; Sequence 3, Application PC/TUS0527579
; GENERAL INFORMATION:
; APPLICANT: Mount Sinai School of Medicine
; APPLICANT: Ioannou, Yiannis
; APPLICANT: Davies, Joanna P.
; TITLE OF INVENTION: NPC1L1 AND NPC1L1 INHIBITORS AND METHODS OF USE THEREOF
; FILE REFERENCE: 2201581-WOO
; CURRENT APPLICATION NUMBER: PCT/US05/27579
; CURRENT FILING DATE: 2005-08-12
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US05-27579-3

Query Match 78.3%; Score 5407; DB 1; Length 1333;

Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy 4 AGLRGWLLWALLLELAQSEPYTTHQPGYCAFYDECGKPELGSGLMTLSNVSCLNTPA 63
Db 3 AAQGWLLWALLLNSAQGLYPTTHKAGFCTFYBECGKPELGGTSLTSLNVSCLNTPA 62
Qy 64 RKITGDHLILLQKICPRLYTGN-TOACCSAKVLSLEASLSITKALLTRCPACSNFVN 122
Db 63 RHVTGDHALLQVCPRLYNGPNDYVACCTKQLVSLDSSLITKALLTRCPACSNFVS 122
Qy 123 LHCHNTCSNQSILFINVTRVAQGLPVAVYEAFOHSPAEQSYDSCSRVVRPAAT 182
Db 123 IHCHNTCSQDSILFINVTRVQDPCQLPAVYEAFOHSPAEQSYDSCSRVVRPAAS 182
Qy 183 LAVGTMCVYGSALCNAORWLNFGQDTGNGLAPLDTTHLLRPGQAVGSGIQPLNRSVAR 242
Db 183 LAVGSMCVYGSALCNAORWLNFGQDTGNGLAPLDTTHLLRPGQALADGMKPLDKGKITP 242
Qy 243 CNESQDDVATCSQDCAASCPAIAARPQALDSTFYLGOMPGSLVLIILCSVFPAVTIL 302
Db 243 CNESQGEDSAACSQDCAASCPVIPPALRPSFYMGPMGLALIIITAVFVLSVVL 302
Qy 303 VGRVAPARDKSMVDPPKGTSLSDKLSPTHTLLGQFPQGMGTWVASMPLTILVLSVIP 362
Db 303 VYLRVASNRKNKTAGSQEAPNLPRKRPSPTVLGRPFESWGTRVASMPLTVLALSFI 362
Qy 363 VVALAAGLVPTLTTPDVELMSAPNSQARSEKAFHDOHEGPPRTNOVILTAPNRSYRY 422
Db 363 VIALSVGLTFIETLTPDVELMSAPNSQARSEKAFHDOHEGPPRTNOVILTAPNRSYRY 422
Qy 423 DSLLLGPKNFGILDLLELELELERLHLQVMSPEAQRNLSLQDICIAPLNPDNTSL 482
Db 423 DSLLLGPKNFGILDLLELELELERLHLQVMSPEAQRNLSLQDICIAPLNPDNTSL 482
Qy 483 YDCCINSLLQYFONNRTLLLTANQTLMGQTSQVMDKHFLYCANAPLTKDGTALALSC 542
Db 483 TDCCVNSLLQYFONNHTLLLTANQTLMGQTSQVMDKHFLYCANAPLTKDGTALALSC 542
Qy 543 MADIYAPVPPFLAIGYKGDYSEBALIMTSLNNYPAGDPRLAQAKLWEERAFLEEMRA 602
Db 543 IADYAPVPPFLAVGQYQGTDYSEBALIITFSINYPADDPMAHAKLWEERAFLEEMOS 602
Qy 603 FORRMAGMFQVTFATERSLEDEINRTTAEPLIATSYIVIFLYISIALGSGSRWRSVAV 662
Db 603 FORSTADKQIARSAERSLEDEINRTTIQDLPVFAISYLIVFYISIALGSGSRWRSVAV 662
Qy 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGIRSSILVILQVVPFLVSVGADNIFIFVLEYQ 722
Db 663 DSKATLGLGVAVVLGAVMAAMGFFSYLGIRSSILVILQVVPFLVSVGADNIFIFVLEYQ 722
Qy 723 RLPRRGEPREVHIGALGRVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMPEGEAREHIGRTLGSVAPSMLLCSLSEAI CFFLGALTPMPAVRTFALTSGLAIVL 782
Qy 783 DFLQMSAFVALLSLDSKQEARSLDVCCVCKPQELPPQCGEGLLGFKYAFKAYAPFLH 842
Db 783 DFLQMTAFVALLSLDSKQEARSDPVCCFSSRLNPPPKQKGLLCCFRKIYTFPLH 842
Qy 843 WITRGVLLFLFALFGVSYLSMCHISVGLDQELALPKDSYLDYFYLFLNRYFEVGPVVF 902
Db 843 RPIRPVLLFLVLFGANLYLMCNISVGLDQDLALPKDSYLDYFYLFLNRYFEVGPVVF 902
Qy 903 VTLGYNFSSAGMNAICSSAGCNFSPFTQKIYATEPPEQSYLAIPASSWVDDFIDWLT 962
Db 903 DTTSGYNFSTEAGMNAICSSAGCESFSLTQKIYASEFPNQSYVAIAASSWVDDFIDWLT 962
Qy 963 P-SSCCRLYISGPNKDKFCPTVNSLNCNKMSITWGSVRSPVQFHKYLPWFLNDRN 1021
Db 963 PSSCCRIYTRGPHKDEFCEPSTDTSFNCLKNMRTLGPRVPTTEQFHKYLPWFLNDRN 1022
Qy 1022 IKCPKGLAAYSTVNLTSVGQVLAASRMAHYKPKNSQDYTEALRAARELANITADLR 1081
Db 1023 IKCPKGLAAYSTVNLSDGQIIASQFMAHYKPLNSQDFTEALRASRLANITAEUR 1082

Db 1023 IRCPKGGLAAVYRTSVNLSSDQGIASQFMAVYHKPLRNSQDFTREALRASLLAANITAEIR 1082
Qy 1082 KVPGETDPAFEVFPYITITNVFYEQYITILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLN 1141
Db 1083 KVPGETDPAFEVFPYITITNVFYEQYITILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLN 1142
Qy 1142 LLSIVMILVDTVGFMAWDISYNAVSLNLYSANGMSVEFVSHITRSPAIKTPWLEA 1201
Db 1143 LLSIVMILVDTVGFMAWDISYNAVSLNLYSANGMSVEFVSHITRSPAIKTPWLEA 1202
Qy 1202 KEATISMSGSAVAGVAMTNLPGILVLGLAKAQIQLIFFFRNLNLTLLGLLHGLVFLPV 1261
Db 1262 KDATIFMGSAGVAGVAMTNLPGILVLGLAKAQIQLIFFFRNLNLTLLGLLHGLVFLPV 1262
Qy 1262 LSYGPDVNPALAEQKRAEAAVAVASCPNHPFSRVSTADNIYVNHSPFGS-1KGAGA 1320
Db 1263 LSYGPDVNPALAEQKRAEAAVAVASCPNHPFSRVSTADNIYVNHSPFGS-1KGAGA 1321
Qy 1321 ISNPLPNNGRQF 1332
Db 1322 ASSSLPKSQDKF 1333

RESULT 46

US-10-621-758A-12

; Sequence 12, Application US/10621758A

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JBO1603-K-US

; CURRENT APPLICATION NUMBER: US/10/621,758A

; PRIORITY FILING DATE: 2003-07-17

; PRIORITY FILING DATE: 2002-07-19

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: Patent in version 3.1

; SEQ ID NO 12

; LENGTH: 1333

; TYPE: PRT

; ORGANISM: Mus sp.

US-10-621-758A-12

Query Match

Best Local Similarity 77.2%; Score 5407; DB 36; Length 1333;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy 4 AGURGLWALLRLAQSEPTTTHQPGYCAFDGCKNPELSGSLMTLSNVCSLSNTPA 63
Db 3 AAQOGLWALLLSAQSGLYPTTHKAGCTFYBCKRNPESLGLTSLNSISCSNTPA 62
Qy 64 RKITGDHLILLOKICPLRYTGN-TOACCSAQVLSLEASLITKALLTRCPACSDNFVN 122
Db 63 RHVTGDHALLORVCPRLYNGNDYACCSLQVLSLSSITKALLTRCPACSDNFVN 122
Qy 123 LHCHNTCSNQLSFINVTRVQGLGQLPVAVAYEAFYQHSFABQSDYSCSRVVPAAAT 182
Db 123 LHCHNTCSNQLSFINVTRVQGLGQLPVAVAYEAFYQHSFABQSDYSCSRVVPAAAT 182
Qy 183 LAVTGCMGYGSAICNAQWILNFGQDTGNGLAPLDTFHLBPQGVGSGIOPANEVAV 242
Db 183 LAVTGCMGYGSAICNAQWILNFGQDTGNGLAPLDTFHLBPQGVGSGIOPANEVAV 242
Qy 243 CNESQDDVATCSQDCAASCPAIPARQALDSTFFYLGOMPGSLVLIILCSVFVAVTILL 302
Db 243 CNESQDDVATCSQDCAASCPVIPPALRPSFYGRMPGWLALIIITAVVLLSVL 302
Qy 303 VGFVRVAPDKSKMDPKKGTSLSDKLSFSTHTLQGFQGGTGWASWPLTILVLSVTP 362
Db 303 VYLVRVASNKNKTAGSQEAPNLPKRKRPSPHTVLGRFFESWGTRVVASWPLTILVLSVTP 362

```

; APPLICANT: Altmann, Scott W
; APPLICANT: Wang, Luquan
; APPLICANT: Graziano, Michael
; APPLICANT: Murgolo, Nick
; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF
; FILE REFERENCE: JB01603-K1-US
; CURRENT APPLICATION NUMBER: US/10/646,301A
; CURRENT FILING DATE: 2003-08-22
; PRIOR APPLICATION NUMBER: 60/397,442
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/621,758
; PRIOR FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus sp.
US-10-646-301A-12

Query Match      78.3%; Score 5407; DB 36; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGWLLALLRLAQSEPTTHIQPGYCAPYDECGKPELSGLMTLSNVSLCNTPA 63
DB 3 AAQGWLLALLNSAQSELYTTHKAGFCFTYEECGKPELSGLTSLNSICLSNTPA 62

QY 64 RKITGPHLLILLOKICPLRLYTGPN-TQACCSAKOLVLEASLSITKALLTRCPACSNFVN 122
DB 63 RHVTGPHALLQVCEPLRYNGPNDTYACCSTQQLVSLDSLSITKALLTRCPACSNFVS 122

QY 123 LHCHNTCSNPQSLFINVTRVAQAGOLPAVAYEAFYQHSFAEQSYDSCSRVVPAAAT 182
DB 123 IHCHNTCSPDQSLFINVTRVQDPQLPAVAYEAFYQHSFAEKAYESCSRVIPAAAS 182

QY 183 LAVGTWCVGYGALCNAQRLWLFQGTGNGLAPLDTTFHLLPQGVAGSGIOPLNQGVAR 242
DB 183 LAVGSMCVGYGALCNAQRLWLFQGTGNGLAPLDTTFHLLPQGVAGSGIOPLNQGVAR 242

QY 243 CNEQGDVVATCSQDCAACPAIARPOALDSTFYLQMGPGSILVLIILCSFPAVVTILL 302
DB 243 CNEQGDVSAACQDCAACSCVIPPALRPSFYNGRMPGMLALIIITFAVFLLSVVL 302

QY 303 VGRVAPARDKSNVDPKKGTSLSLKSPTSHTLLQPFQGMGTWVASHPLTILVLSVIP 362
DB 303 VYLRVASNENKNTAGSQEAPNLPRKRFSPHTVLGRFFESWGTRVASWPLTVLALSFI 362

QY 363 VVALAAGLVFTELTTPDVELWAPNSQARSEKAFHDQHFPPFRTNQVILTAPNRSYRY 422
DB 363 VIALSVGLTFIELTTPDVELWAPNSQARSEKAFHDQHFPPFRTNQVILTAPNRSYRY 422

QY 423 DSILLGPKNFSGLTDLLELLELLELLELLELLELLELLELLELLELLELLELLEL 482
DB 423 DSILLGPKNFSGLTDLLELLELLELLELLELLELLELLELLELLELLELLEL 482

QY 483 YDCCINSLLQYQNNRTLLLTANQTLGQTSQVMDKHFLYCANAPLTFKDGTALALSC 542
DB 483 TDCCVNSLLQYQNNHTLLLTANQTLGQTSQVMDKHFLYCANAPLTFKDGTALALSC 542

QY 543 MADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLLNYPAGDPRLAQAKLWEEAFLEEMRA 602
DB 543 IADYGAPVFPFLAIGYKGYKGYSEABALIMTFSLLNYPAGDPRLAQAKLWEEAFLEEMRA 602

QY 603 FORMAGMFOVTFEATERSLEDEINRTAEDLIPATSYVIFLYTSLALGSSYSSRSRVW 662
DB 603 FORSTADKFOIAFSAERSLEDEINRTTIQDLVPFAISLYLIVLYTSLALGSSYSSRSRV 662

QY 663 DSKATLGLGVAVLGVANMAAGFFSYLGRSLVILQVVPFLVSVGADNIPFVLELYQ 722
DB 663 DSKATLGLGVAVLGVANMAAGFFSYLGRSLVILQVVPFLVSVGADNIPFVLELYQ 722

QY 723 RLPRPGEPREVHIGRALGRVAPSMLLCSLSEAICFFLGALTTPMPAVRTFALTSGIAVL 782

```

RESULT 48

US-10-663-208A-12

; Sequence 12, Application US/10663208A

; GENERAL INFORMATION:

; APPLICANT: Altmann, Scott W

; APPLICANT: Wang, Luquan

; APPLICANT: Graziano, Michael

; APPLICANT: Murgolo, Nick

; TITLE OF INVENTION: NPC111 (NPC3) AND METHODS OF USE THEREOF

; FILE REFERENCE: JB01603K2 US

; CURRENT APPLICATION NUMBER: US/10/663,208A

; CURRENT FILING DATE: 2003-09-16

; PRIOR APPLICATION NUMBER: 60/397,442

; PRIOR FILING DATE: 2002-07-19

; PRIOR APPLICATION NUMBER: 10/621,758

; PRIOR FILING DATE: 2003-07-17

; PRIOR APPLICATION NUMBER: 10/646,301

; PRIOR FILING DATE: 2003-08-22

; NUMBER OF SEQ ID NOS: 50

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 12

; LENGTH: 1333

; TYPE: PRT

; ORGANISM: Mus sp.

US-10-663-208A-12

Query Match

Best Local Similarity 78.3%; Score 5407; DB 36; Length 1333;

Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

Qy	4	AGLRGWLWALLRLAQSEPYTTTHQGYCAFYDECGKNPELSGLMTLSNVSLCNTPA	63
Db	3	AAMQGWLWALLNSAQGELYTPTHKAGCTFYEECGKNPELSGGLTSLNISCLSNTPA	62
Qy	64	RKITGDHLLILOKICPLRYTGN-TQACSAKQIUVSLSEASLSTIKALLTRCPACSNFVN	122
Db	63	RHVTDGHLALQYCPRLYNGPNDTYACSTKQIUVSLDSSLSTIKALLTRCPACSNFVS	122
Qy	123	LHCNHTCSPNOSLFINTRVQAQIAGAGQLPAVAYEAFYQHSFAEOSYDSCSRVVRPAAT	182
Db	123	IHCNHTCSPQSLFINTRVVQDPGQLPAVAYEAFYQHSFAEKAYESSCRVRIIPAAAS	182
Qy	183	LAVTMCGVYGSALCNAQRMLNFQDGTGNGLAPLDITFHLLPEQAVGSGIOPINEGVAR	242
Db	183	LAGSMCGVYGSALCNAQRMLNFQDGTGNGLAPLDITFHLLPEQOALADGMKPLDGKITP	242
Qy	243	CNESQGDVVATCSQDCAASCAPATARQALDSIFYLQMGPSVLIIILCSFANVTILL	302
Db	243	CNESQGSDSAACSCQDCAACPVIPPPALLRPFYMGPMGFWLALIIIFAVFVLLSVWL	302
Qy	303	VGFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPLTILVLSVIP	362
Db	303	VYLRVANRNKNKTAGSQEAPNLPKRKRPSPHTVILGRFPESWGTRVASWPLTVLASFIV	362
Qy	363	VVALAAGLVFTFELTDPVELWSAPNSQARSEKAFHQHFGFPFRTNQVILTAPNRSRYR	422
Db	363	VIALSVGLTFTFELTDPVELWSAPKSQARKEKAFHDEHFGFPFRTNQIFVTAKNRSYKY	422
Qy	423	DSLALLGPNFSGIILDLLELELOERLRLHQLVMSPEAORNTSLODICVAPLNPONTSL	482
Db	423	DSLALLGPNFSGIILSLOLLELOERLRLHQLVMSHEAQRNLSLODICVAPLNPHTNSL	482
Qy	483	YDCINSILQYFQNNRTLLILANTQITLMGQTSQVMDKHFLYCANAPLTFKDGTALALSC	542
Db	483	TDCCVNSILQYFQNNHTLLLILANTQITLNGQTSVMDKHFLYCANAPLTFKDGTALALSC	542
Qy	543	MADYGAPVPPPLATGGYKGKDYSEAEALIMTFSLNNYPAGDPRLAQAUKMEEAFLEMBRA	602
Db	543	IADYGAPVPPPLAVGGYQGTDYSEAEALIIITFSINNPADDPRLMAHAKLWEEAFLEMQS	602
Qy	603	FORRMAGFQVYFTFAERSLEDEINRTTAEDLPIEATSYVIFELVISALGYSYSSWVMV	662
Db	603	FORSTADKQIFAFSAERSLEDEINRTTIQDLVFAISLIVFLVISALGYSYGRWSKAV	662
Qy	663	DSKATLGLGGVAVVLGAVMAAMGFYSYLGIRSSIVILQVVPFLVLSVGADNIIFFVLEYQ	722
Db	663	DSKATLGLGGVAVVLGAVVAAMGFYSYLGVPSSVLIIQVVPFLVLAVGADNIIFFVLEYQ	722
Qy	723	RLPRRPGEPREVHIGRALGRVAPSMKILCSLSEALCFPIGALTMPAVRFTALTSGLAAIL	782
Db	723	RLRMPGEGREAHIGRTLGSVAPSMKILCSLSEALCFPIGALTSPAVRFTALTSGLAIIIF	782
Qy	783	DFLQMSAFVALLSLDSKROEASRLDVCCCKVPQELPPPGQEGELLIGFPQKAVAPFLH	842
Db	783	DFLQMTAFVALLSLDSKROEASRPDVVCCPSSRNLPFPKQEGELLICFFRKIYTPFLH	842
Qy	843	WITRGVLLLFIALFGVSLYSIMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPYVF	902
Db	843	RPIRPVLLLFLVLFGANLYLMWCNISVGLDQDLALPKDSYLDIDYFLFLNRYLEVGPPYVF	902
Qy	903	VTTILGYNPSSBAGNNAICSSAGCNFSFTQIKIYATBEPQSYLAIIPASSWDDPIDMLT	962
Db	903	DTTSGYNFSTBAGNNAICSSAGCSFSLTQIKIYASBFPNQSYVIAIASSWDDPIDMLT	962
Qy	963	P-SSCCRLYISGPNKDKFCPTVNSLNCNKMCSITMGSVRAPSVQEOPHKYLPWFLNDRPN	1021
Db	963	PSSCCRLYITGPHKDFECPTDTSFNCLKNCMRTILGPVPTTEQHFHYLPWFLNDTPN	1022
Qy	1022	IKCPKGGIAAYSTSVNLTSQGVYASRPMAYHKPLKNSQDYTEALRAARELAANITADLR	1081
Db	1023	IRCPKGGIAAYRTSVNLSSDQIITASQPMAYHKPLRNSQDTEALRASRLIAANITAELR	1082

RESULT 49

```

US-10-736-769-12
: Sequence 12, Application US/10736769
: GENERAL INFORMATION:
: APPLICANT: Altmann, Scott W
: APPLICANT: Wang, Luquan
: APPLICANT: Graziano, Michael
: APPLICANT: Murgolo, Nick
: TITLE OF INVENTION: NPC11 (NPC3) AND METHODS OF USE THEREOF
: FILE REFERENCE: J01603-K3-US
: CURRENT APPLICATION NUMBER: US/10/736,769
: CURRENT FILING DATE: 2003-12-16
: PRIOR APPLICATION NUMBER: 60/397,442
: PRIOR FILING DATE: 2002-07-19
: PRIOR APPLICATION NUMBER: 10/621,758
: PRIOR FILING DATE: 2003-07-17
: PRIOR APPLICATION NUMBER: 10/646,301
: PRIOR FILING DATE: 2003-08-22
: PRIOR APPLICATION NUMBER: 10/663,208
: PRIOR FILING DATE: 2003-09-16
: NUMBER OF SEQ ID NOS: 51
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 12
: LENGTH: 1333
: TYPE: PRT
: ORGANISM: Mus sp.
US-10-736-769-12

```

Query Match	78.3%;	Score 5407;	DB 37;	Length 1333;
Best Local Similarity	77.2%;	Pred. No. 0;		
Matches 1028;	Conservative 126;	Mismatches 174;	Indels 4;	Gaps 4;
QY	4	AGLRGWLIIWALLRLAQSEBYTHIQBGYCAFYDECGKNPGLSGISLMTLSNVCSLSNTPA	63	
DB	3	AAMQGWLLWALLNSAQGLYTPHKGAGCTFYEECKGNPGLSGGLTSLNISCSLSNTPA	62	
QY	64	RKITGDHLILLOKICPLRYLTPGN-TOACCSAKQLVSLSEASLSITKALLTRCPACSDNFVN	122	
DB	63	RHVDTGSHALLQORVCPRLYNGPNDTYACCTSKQLVSLDSSLSTIKALLTRCPACSENFVS	122	
QY	123	LHCHNTCSPNQSLFINVTRVAQLGAGQLPAVAYEAFYQHSFAEQSYDCSRVRVPAAAAT	182	
DB	123	IHCHNTCSPQSLFINVTRVQRDPGQLPAVAYEAFYQHSFAEQSRSRVRIPAAS	182	
QY	183	LAVTMCGVYGSALCNARQWLNFGDGTGNGLAPLDITFHLLEPQAVGSGIQLPNEGVAR	242	
DB	183	LAVSGMCGVYGSALCNARQWLNFGDGTGNGLAPLDITFHLLEPQALADGMPKPLDGKITP	242	
QY	243	CNESQGDVVATCSQDCAASCPAIARQALDSTFVLGMPGCSVLIIILCSVFPAVVTILL	302	
DB	243	CNESQGDSENAACSQDCAASCPIVPPPPALRPSFYMGMPGWLALIIIFVAVFVLLSVVL	302	


```
Db 473 PPLDIQILHQLVDLQIAIEN--ITASYDNETVTLQDICIAPLSPYNT---NCTILSVLNY 527
Qy 494 FQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKQGTALALSCMADYGAPVPPF 553
Db 528 FQNSHVLVDHKKGGDDFF---VYADYTHFLYCVRAPASLNDTSLDPCLTGTFGGFVFPW 584
Qy 554 LAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFOV 613
Db 585 LVLGYYDQYNNATLVITPPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTI 641
Qy 614 TPTAERSLEDEINRTAEDLPIFATSYIVIFLYISALGSYSSWRVMVDSKATLGLGV 673
Db 642 SFTAERSIEDELNRESDSVFTVVIYAIMFLYISALGHKSCRRLLVDSKVLGIAGI 701
Qy 674 AVVLGAVMAAMGFYSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRGEPRE 733
Db 702 LIVLSSVACSLGVFSYIGLPLTVIEVIFLVAVGVNDNIFILVQAYORDERLQGETLD 761
Qy 734 VHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSGLAVILDFFLQMSAFVA 793
Db 762 QQLGRVLGEVAPSMFLSFSFSEIVAFGLNLSVMPAVHTFSLFAGLAVFDLLOITCFVS 821
Qy 794 LLSLDSKQKEASRLDVCCVKPQELPPGQ-CEGILLGFFQKAYAPFLLHMTIRGVLL 852
Db 822 LLGLDIKQKRNRLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLLKQWRPIVIAI 881
Qy 853 FLALFGVLSYNSCHTISVGLDQELALPKDSYLLDYFLFNRYEVCAPYFVTLGYNFS 912
Db 882 FVGVLFSFSAVLNKVDIGLDQSLSPDDSYMDYFKSISQYLHAGPPYFVFLBEGHDYTS 941
Qy 913 EAGMNAICSSAGCNSFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDMLTP--SSCCRLYI 971
Db 942 SKGQNVCGGCGNDSLVQIIFNAQOLDNYTRIGFAPSSWIDDFYDWKQSSCCRV-- 999
Qy 972 SGPNKDKFCPSVTNSLNCNCKMT-MGSVRPSVEQPHYLPLWFLNDRPNITCKPKGLA 1030
Db 1000 -DNITDQFCNASVDPACVR-CRPLTPEGKRPQGGDFMRFLPMLSDNPNPKCGKGHA 1057
Qy 1031 AVYSTN--LTSQGVLASREMAHKPLKNSODYTEALRAARELANITADLRKVPGETD 1088
Db 1058 AVSSAVNLLGHGRTVGATFYTHVTLQTSADFIDALKARLIASNV-TWGMINGS-- 1114
Qy 1089 AFEVFPYTTINVFYEQYLTILPEGLFMLSCLVPTFAVSCLLLGLDLSGLNLLSIYMI 1148
Db 1115 ARVFPYSYVFYFYEQLTIIDTIFNLGVSILGAIFLVTWLLGCELSAVIMCATIAMV 1174
Qy 1149 LVDTVGFMAWLDISYNAVSLINLVSAGVMSVFVSHITRSPAITKPTWLERAKATISM 1208
Db 1175 LVNMFQVMWLGISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKGSVERAEBALAHM 1234
Qy 1209 GSAVAGVAMTWLPGTLVLGLAKAQLIQIFPERLNLITLLGLLHCLVPLVLSVGPD 1268
Db 1235 GSSVFSGITLTKFGGIWLAFAKSQIFQIFRYMVLAMVLLGHTLPLPLVLLSYIGPS 1294
Qy 1269 VNPA 1272
Db 1295 VNKA 1298
```

RESULT 54

US-10-940-774A-10152

; Sequence 10152, Application US/10940774A

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/10/940,774A

; CURRENT FILING DATE: 2004-09-15

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

```
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10152
; LENGTH: 1318
; TYPE: PRT
; ORGANISM: Human
US-10-940-774A-10152
```

```
Query Match          34.8%; Score 2403; DB 39; Length 1318;
Best Local Similarity 39.7%; Pred. No. 2.2e-201;
Matches 525; Conservative 238; Mismatches 451; Indels 110; Gaps 26;
```

```
Qy 2 AEAGLRGWLLWA-----LLRLAQSEPYTIHQPGYCAFYDECG-----KNPELSG 47
Db 32 ASALRGHSMTARGIALGLGLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEYSG 86
Qy 48 SLMTLSNVCSLSTNPARKITGDHLITLQKICPRLYTGPNTQACCSAKQLVLSLEASLITK 107
Db 87 -----PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPL 131
Qy 108 ALLTRCPACSNFVNLHCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAEYAFYQHS 163
Db 132 QFLSRCPSCFYNNLNLFCELTCSPRSQFLNVTATEDYDVPVTNQTKTNKELQYVVGOS 191
Qy 164 FAEQSYSCSRVRPAAATLAVGTMCYVYGSAICNAQRLNFPQDGTGNGLAPLDIT---- 219
Db 192 FANAMYNACRDEAPSSNDKALGLLCKGADA-CNATWIEYFNKNDQGAQPTTITPVS 250
Qy 220 ---FHLLEFGQVSGIQIPLNEGVARCNESQGDVATCQDCAASCAPAIARPQ----- 270
Db 251 DFPVH-----GMEPMNATKGCDESDEVTAPCSCDCSIVCGPKQPQPPPPAPW 300
Qy 271 ---ALDSTFYLGQWPGSLVLIILCSFPAVVTILL-----VGRVAPARDKS 314
Db 301 TILGLDAMVIMWITWYMAFLVFFGAFPAVMCYRKRYFVSEYTPIDSNIAFSV-NASDKG 359
Qy 315 KMVDPKGTSLSDKLSFSTHTLLGQFFQGWGTWVASWPLTILVLVSVIPVVALAAGLVFTE 374
Db 360 E-----ASCCDPVSAAFEGCLRLPFRWGSFCVRNPGCVIFFSLVITACSSGLVFR 412
Qy 375 LTTPDVELMSAPNSQARSEKAFHDQHPFPFRTNQVILTAPNRSYRYDSLLLGPK-NPS 433
Db 413 VTTNPVDLMSAPSSQARLEKEYFDQHPGFFRTEQLIIRAPLTDKHIQYPSGADVPRG 472
Qy 434 GILDLDLLELLELQERLHRHQVMSPEAQRNLSQDICVAPLNPDNTSLDYDCINSLLQY 493
Db 473 PPLDIQILHQLVDLQIAIEN--ITASYDNETVTLQDICIAPLSPYNT---NCTILSVLNY 527
Qy 494 FQNNRTLLLTANQTLMGQTSQVMDKDHFLYCANAPLTFKQGTALALSCMADYGAPVPPF 553
Db 528 FQNSHVLVDHKKGGDDFF---VYADYTHFLYCVRAPASLNDTSLDPCLTGTFGGFVFPW 584
Qy 554 LAIGYKGYKDYSEAEALIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMFOV 613
Db 585 LVLGYYDQYNNATLVITPPVNNYNDTEKLQRAQAWKEFINFVKYKN---PNLTI 641
Qy 614 TPTAERSLEDEINRTAEDLPIFATSYIVIFLYISALGSYSSWRVMVDSKATLGLGV 673
Db 642 SFTAERSIEDELNRESDSVFTVVIYAIMFLYISALGHKSCRRLLVDSKVLGIAGI 701
Qy 674 AVVLGAVMAAMGFYSYLGIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRGEPRE 733
Db 702 LIVLSSVACSLGVFSYIGLPLTVIEVIFLVAVGVNDNIFILVQAYORDERLQGETLD 761
Qy 734 VHIGRALGRVAPSMLLCSLSEACFFLGALTMPAVRTFALTSGLAVILDFFLQMSAFVA 793
Db 762 QQLGRVLGEVAPSMFLSFSFSEIVAFGLNLSVMPAVHTFSLFAGLAVFDLLOITCFVS 821
Qy 794 LLSLDSKQKEASRLDVCCVKPQELPPGQ-CEGILLGFFQKAYAPFLLHMTIRGVLL 852
Db 822 LLGLDIKQKRNRLDIFCCVGAEDGTSVQASESCLFRFFKNSYSPLLLKQWRPIVIAI 881
```



```
PCT-US03-40978-1542
; Sequence 1542, Application PC/TUS0340978
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: PCT/US03/40978
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-40978-1542

Query Match      34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy      7  RGWLWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
Db      4  RGLALGLLLLLCPAQVFSQ-----SCVMYGCGLIAYGDKRYNCEYSG-----46

Qy      60  NTPARKITGDHLLILKQICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db      47  --PPKPLPKDGVLDVQELCPGFFFG-NVSLCCDVRQLQTLKDNQLPLQPLQSRCSFCFYN 103

Qy      120  FVNHLCHNTCSNQSLFINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db      104  LNLFCELTCSRQSOFLNVTATEDYDVPVNTQKTNVKELYYQGSFANAMYNACRDV 163

Qy      176  RVPAAATLAVGTWCMGYGSGALCNAQRWLNFGQDTGNGLAPLDIT-----PHLLPQGA 228
Db      164  BAPSNKALGLGCKDADA-CNATNWIEMFNKNGQAPFTITPVFSDFPVH-----215

Qy      229  VCSGTQPLNEGVARCNESGDDVATCSQDCAASCAPATAPQ-----ALDSTFYLG 279
Db      216  ---GMEPMNNAATKGDCEVDEVTAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272

Qy      280  OMPGSLVLIILCSVPVAVTILL-----VGRVPAPARDKSMVDPKKGTSL 326
Db      273  WITYWAFLLVPGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy      327  DKLSFSTHTLLGQFGGWTWASWPLTLLVLSVPVVALAAGLVFTTELTPDVELWSAP 386
Db      325  DPVSAAFEGCLRLFRMGSCFVRNPGCVIFFSLVFTACSSGLVFRVTTNPVDLWSAP 384

Qy      387  NSQARSEKAFDQHGPPFRTNQVILTAPNRSYRVDLSLLGPK-NFSGILDLDLLELL 445
Db      385  SSOARLEKEYFDHGFPGFFRTTQELIRAPLTKHIYQYPYSGADVPFGPPLDIQLHQVL 444

Qy      446  ELQERLRHLQVMSPEAQRNISLDIYCAPLNPDNTSLYDCINSLLIQFQNNRLLLLTA 505
Db      445  DLQIAIEN--ITASYDNEVTIQLDCLAPLSYNT---NCTLSLVNLYQNSHSLDHKK 499

Qy      506  NOTLMGQTSQVDKDHFLYCANAPITFKDGTALALSCMADYGAQVPPFLAIGYKGKDY 565
Db      500  GDDFF---VYADVHTFLYCVRAPASLNDTSLLDHDCILGTFGPGVPPMLVLGDDQNYN 556

Qy      566  EBAALIMTSLNYPAGDPRLAQAKLWEAFLEEMRAPQRMAGMPQVTFTAERSLEDEI 625
Db      557  NATALVITFPVNNYNDTEKLQRAQAWKEFTNFVKNYKN---PNUTISFTAERSIEDBL 613

Qy      626  NRTTARDLPIFATSYIVIFLYTSLALGYSNRSVMVDKATLGLGGAVALVGNVMAAG 685
Db      614  NRESDSVITVVISYAIMELYSLALCHTKSCRRLLVDSKVSGLGIALIIVSSVACSLG 673

Qy      686  FFSYLGISSVLQVVPFLVLSVGNADNIFIFVLEYQRLPRPPEPREVHI GRALGRVAP 745
Db      674  VFSYIGLTLTIVIEVIFLVLAVGVNDNIFILVOAYORDERLQGETLDQQLGRVLGEVAP 733
```

```
Qy      746  SMLLCSISEAICFFLIGALTMPAVRTFALTSGLAVILDQLQMSAFVALLSLSRQREAS 805
Db      734  SMFLSSFSFETVAFFLGALSVMVAHTFSLFAGLAVFIDFLAQITCFVSLGLDIKQREKN 793

Qy      806  RLDVCCCVKPOQLPPPGQ--GGLLLGFPQKAYAPLLHWITRGVVLVLLFLALFGVLSYGM 864
Db      794  RLDIFCCVRGAEDGTSTQASESCLFRFPFKNSYSPILLKDMWRPIVIAIFVGVLFSFI 853

Qy      865  CHISVGLDOELALPKDSYLLDVELFLNRYFEVGAPVYFVTTLYGNFSEAGMNAICSSAG 924
Db      854  NKVDIGLDQSLSDMPDDSYMVDYFKSISQYLHAGPPVYFVLEEGHDYTSKQGNMVCYGM 913

Qy      925  CNPFSTQIKIYATEFEPEQSYLAIPASSWVDDFDLWLP--SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDLSLVQOIFNAQAQLDNYTRIGFAPSSWIDDYDPMVKPQSSCCRV---DNITDQFCNAS 970

Qy      984  VNSNCLKNCMSIT--MGSVRPSVQFHKYLPWFLNDRNIKCPKGLLAAYSTSVN--LTS 1040
Db      971  VVDPACVR--CRPLTPEGKRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNILLGH 1029

Qy      1041  DGQVLASFEMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYTTNV 1100
Db      1030  GTRVATGYFTMYHTVLTQTSADFIDALKKARLIASNVT--ETWINGS--AYRVFPYSVFY 1086

Qy      1101  FVEQYLTILPGLFMLSCLVPTFAVSCLLGLDLRSGLLNLSIVMLIVDTVGFMAWD 1160
Db      1087  FVEQYLTIIDTIFNLGVSIGAIPLVTVLWGLCELWSAVIMCATIAMVLVNMFGVWLWG 1146

Qy      1161  ISYNAVSLINLVAVGMSVFEVSHITRSPASTKTWLERAKEATISMGSAVPAVAMTN 1220
Db      1147  ISLANVSLVNLVMSGSGISVEFCSHITRAFTVSMKGSVEREAEALAHMGSSVFSGITLTK 1206

Qy      1221  LPGILVILGLAKAQLIQIFERLNLILLGLLGLVLELVPVILSYVGGPDVNP 1272
Db      1207  FGIIVVTLAPAKSQIPQIFIFRYMLANVLLGATHGLIFLPLVLSYIGPSVNKA 1258

RESULT 57
PCT-US04-07141-810
; Sequence 810, Application PC/TUS0407141
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: PCT/US04/07141
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 810
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US04-07141-810

Query Match      34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy      7  RGWLWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
Db      4  RGLALGLLLLLCPAQVFSQ-----SCVMYGCGLIAYGDKRYNCEYSG-----46

Qy      60  NTPARKITGDHLLILKQICPRLYTGPNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db      47  --PPKPLPKDGVLDVQELCPGFFFG-NVSLCCDVRQLQTLKDNQLPLQPLQSRCSFCFYN 103

Qy      120  FVNHLCHNTCSNQSLFINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db      104  LNLFCELTCSRQSOFLNVTATEDYDVPVNTQKTNVKELYYQGSFANAMYNACRDV 163

Qy      176  RVPAAATLAVGTWCMGYGSGALCNAQRWLNFGQDTGNGLAPLDIT-----PHLLPQGA 228
Db      164  BAPSNKALGLGCKDADA-CNATNWIEMFNKNGQAPFTITPVFSDFPVH-----215

Qy      229  VCSGTQPLNEGVARCNESGDDVATCSQDCAASCAPATAPQ-----ALDSTFYLG 279
Db      216  ---GMEPMNNAATKGDCEVDEVTAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272

Qy      280  OMPGSLVLIILCSVPVAVTILL-----VGRVPAPARDKSMVDPKKGTSL 326
Db      273  WITYWAFLLVPGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324

Qy      327  DKLSFSTHTLLGQFGGWTWASWPLTLLVLSVPVVALAAGLVFTTELTPDVELWSAP 386
Db      325  DPVSAAFEGCLRLFRMGSCFVRNPGCVIFFSLVFTACSSGLVFRVTTNPVDLWSAP 384

Qy      387  NSQARSEKAFDQHGPPFRTNQVILTAPNRSYRVDLSLLGPK-NFSGILDLDLLELL 445
Db      385  SSOARLEKEYFDHGFPGFFRTTQELIRAPLTKHIYQYPYSGADVPFGPPLDIQLHQVL 444

Qy      446  ELQERLRHLQVMSPEAQRNISLDIYCAPLNPDNTSLYDCINSLLIQFQNNRLLLLTA 505
Db      445  DLQIAIEN--ITASYDNEVTIQLDCLAPLSYNT---NCTLSLVNLYQNSHSLDHKK 499

Qy      506  NOTLMGQTSQVDKDHFLYCANAPITFKDGTALALSCMADYGAQVPPFLAIGYKGKDY 565
Db      500  GDDFF---VYADVHTFLYCVRAPASLNDTSLLDHDCILGTFGPGVPPMLVLGDDQNYN 556

Qy      566  EBAALIMTSLNYPAGDPRLAQAKLWEAFLEEMRAPQRMAGMPQVTFTAERSLEDEI 625
Db      557  NATALVITFPVNNYNDTEKLQRAQAWKEFTNFVKNYKN---PNUTISFTAERSIEDBL 613

Qy      626  NRTTARDLPIFATSYIVIFLYTSLALGYSNRSVMVDKATLGLGGAVALVGNVMAAG 685
Db      614  NRESDSVITVVISYAIMELYSLALCHTKSCRRLLVDSKVSGLGIALIIVSSVACSLG 673

Qy      686  FFSYLGISSVLQVVPFLVLSVGNADNIFIFVLEYQRLPRPPEPREVHI GRALGRVAP 745
Db      674  VFSYIGLTLTIVIEVIFLVLAVGVNDNIFILVOAYORDERLQGETLDQQLGRVLGEVAP 733
```



```
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVSPDPVH----- 215
Qy 229 VSGIOPLNEGVARCNESQGDVATCSCQCAASCPAIAAPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSCQDCSIVCGPKQPQPPPPAPWPTILGLDAMVIM 272
Qy 280 QMPGSLVLIILCSFPAVVTILL-----VGFVAPADKSKMVDPKGTLSLS 326
Db 273 WITTMALFLVFFGAPFAVMCYRKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSVIPWVALAAGLVFTELTTDPVLSWAP 386
Db 325 DPVSAAPFEGCLRLFTRMGSPCVNRNPGCVIFPSLVPITACSSGLVFRVTNPNVDLWSAP 384
Qy 387 NSQARSEKAFHDQHPGFPFRTRNQVILTAPNRSRYRDSLLLGPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEYFDQHPGFPFRTEQLIIRAPLTDKHIQYPYPSGADVPFGPPDLQILHQL 444
Qy 446 ELQERLRHLQVSPQAQRNLSLODICVAPLNPNTSLYDCCINSLLQYFQNNRTILLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLODICLAPLSYNT---NCTILSVLNYPQNSHVLHDK 499
Qy 506 NOTLMGQTSQVMDKDHFLYCANAPITFKDGTALALSCMADYGAPVFPFLAIGYKGYK 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLHDPCLGTGFGVFPFWLVGLDYDQNYN 556
Qy 566 EABALIMTFSLNYPAGDPRLAQAKLWEBAFLIEMRAFORRMAGMFQVFTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAWEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTADLPIFATSYIVIFLYISLALGSYSSWRVMDSKATIGLGAVVAVLVGAVMAWG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHISKCRLLVDSKVSIGIAGILIVLSSVACSLG 673
Qy 686 FPSYLGRSLVLOVVPFLVLSVGDADNIFIVLEYQRLPRRPEPREVHIGBALGRVAP 745
Db 674 VFSYIGLPLTIVIEVPLVLVAGVDNIFILVQAYQERLQGETLDQQLGRVLGEVAP 733
Qy 746 SMLLCSLSBAICPFLGALTPMAVRTFALTSLGLAVILDLLQMSAFVALISLDSKQREAS 805
Db 734 SMFLSSFSEVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLLDIKQEK 793
Qy 806 RLDVCCVQKQELPPQEG-GEGLLLGFQKAVAPFLLHMITRGVULLFLALFGVSLYSM 864
Db 794 RLDIFCVRGAEDGTSQVASESCLFPEFFKNSYSPLLLDKWMRPIVIAIFGVLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYLFLNRYFEVGAUVYVTTILGYNFSSSEAGNAICSSAG 924
Db 854 NKVDIGLQSLSPDDSYMDYIPKSIISQYLHAGPPVYFVLEEGHDYTSKGQNMVCGMG 913
Qy 925 CNNFSTQKIQVATEPPEQSYLEIPASSWDDPIDMLTP-SSCCRLYISGPNKDKFCPST 983
Db 914 CNNDLSVQIIFNAQLDNYTRIGFAPSSWDYDFDWKVPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT-MGSRVPSVEQPHKLPWFLNDRPNIKCPKGGLAAYSTSVN--LFS 1040
Db 971 VVDPACVR-CRPLTPREGKQRPQGGDFMRFLPMFLSDNPNPKCGKGHAAYSSAVNLLGH 1029
Qy 1041 DQOVLASRWMAHKKPLKNSQDYTEALRAARELANITADLRKVPGTDPAFEVPPYITIV 1100
Db 1030 GTRVGATYPMYTHVTLQTSADFIDALKARLTASNTV-ETMGIINGS--AYRVPYPSVYV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMILVDVTGFMALWD 1160
Db 1087 FREQYLTITDDYIFNLGVSLGAFLVTWVLGCELWSAVIMCMATIAMLVNFMGVNWLWG 1146
Qy 1161 ISYNAVSLINSAVGMVSVEFVSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVRERAREALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVGLAKAQLQIQIFFRNLNLLITLGLHGLVFLPVILSYGPDVNP 1272
Db 1207 FGGIVVLAPAKSQIQIIFYFRMYLAMVLIGATHGLIFLFPVLSYIGPSVYKA 1258
```

RESULT 58

```
PCT-US05-18850-1032
; Sequence 1032, Application PC/TUS0518850
; GENERAL INFORMATION:
; APPLICANT: CHIN, Lynda
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: DPN-059-2
; CURRENT APPLICATION NUMBER: PCT/US05/18850
; CURRENT FILING DATE: 2005-05-27
; NUMBER OF SEQ ID NOS: 1528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1032
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US05-18850-1032
```

```
Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
Qy 7 RGWLLWALLRLAQSEPYTHIHQGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGEGLYAGDKRYNCEYSG----- 46
Qy 60 NTPARKITGDHLLILQKICPRLYTGNTQACCSAKQOLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKONLQLPLOQLSRPCSFYN 103
Qy 120 FVNLHCHNTCSNOSLFINVTR-----VAQIGAGQLPAVVAYEAFYOHSPAESQYSDCSR 175
Db 104 LNLFPCLTCSRQSQFLNVATATEDYVDPVTQTKNVKELQYVYGQSFANAMYNACRDV 163
Qy 176 RVPAATLAVGTCMGYVGSALCNAQRLNFGQDGTGNGLAPLDIT-----PHLLSPQQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVSPDPVH----- 215
Qy 229 VSGIOPLNEGVARCNESQGDVATCSCQCAASCPAIAAPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSCQDCSIVCGPKQPQPPPPAPWPTILGLDAMVIM 272
Qy 280 QMPGSLVLIILCSFPAVVTILL-----VGFVAPADKSKMVDPKGTLSLS 326
Db 273 WITTMALFLVFFGAPFAVMCYRKRYFVSEVTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSVIPWVALAAGLVFTELTTDPVLSWAP 386
Db 325 DPVSAAPFEGCLRLFTRMGSPCVNRNPGCVIFPSLVPITACSSGLVFRVTNPNVDLWSAP 384
Qy 387 NSQARSEKAFHDQHPGFPFRTRNQVILTAPNRSRYRDSLLLGPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEYFDQHPGFPFRTEQLIIRAPLTDKHIQYPYPSGADVPFGPPDLQILHQL 444
Qy 446 ELQERLRHLQVSPQAQRNLSLODICVAPLNPNTSLYDCCINSLLQYFQNNRTILLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLODICLAPLSYNT---NCTILSVLNYPQNSHVLHDK 499
Qy 506 NOTLMGQTSQVMDKDHFLYCANAPITFKDGTALALSCMADYGAPVFPFLAIGYKGYK 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLHDPCLGTGFGVFPFWLVGLDYDQNYN 556
Qy 566 EABALIMTFSLNYPAGDPRLAQAKLWEBAFLIEMRAFORRMAGMFQVFTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAWEKEFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTADLPIFATSYIVIFLYISLALGSYSSWRVMDSKATIGLGAVVAVLVGAVMAWG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHISKCRLLVDSKVSIGIAGILIVLSSVACSLG 673
```

```

686 FFSYLGRSSVLVQVFPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHIGRALGRVAP 745
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
674 VFSYIGLPLTLIVIRVIFPFLVAVGVDNIFILVQAYORDERLOQETLDQOLGRVLGEVAP 733
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
746 SMLCSISEATCFEFGALTPMPAVTFTALTSGLAVIDLFLQWSAFVALLSDLSKQOEAS 805
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
734 SMFLSSFSETVAFFLGLSVMPAVHTFSLFAGLAVFIDFLQLITCFVSLGLDIIKQOEKN 793
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
806 RLDVCCCKVPOELPPPGQ-GEGLLIGFFOKAYAPFLLLHWITRGVLLLLFLALFGVSLYSM 864
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
794 RLDIFCCVRGADGTSVQASCSCLFRFFKNKSVPLLLKDDWMPPIVIAIFGVLSFSIAVL 853
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPYFVTTILGYNFSSEAGMNAICSSAG 924
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
854 NKVDIGLDSLSMPDSDYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKGGQNMVCGMG 913
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
925 CNPFSFTQKIQVATFEPQSYLAIPASSWDDFDIDLTP-SSCCRILYISGPNKDKFCPST 983
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
914 CNNSLVQOQIFNAALQDNYTRIGFAPSSWIDDYFOWVKPQSSCCRV--DNITDQFCNAS 970
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
984 VNSLNLCKNCSIT-MGSVRPSVEGFHKVLPFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
971 VVDPAICVR-CRLPTPEGKQRPQGGDFMRFLPFLSDNPNPKCGKGRAAYSSAVNILLGH 1029
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1041 DGQVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYITNV 1100
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1030 GTRVGATFYMYHTVLTQTSADFIDALKKARLIASNTV-ETMGINGS--AYRVFPYSVFV 1086
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1101 FVEQYLTLPDEGLFMSLCLVPTFAVSCLLGLDLRSLNLLLSIVMLVDTVGFMAWD 1160
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1087 FVEQYLTIIDDTIFNLGVSGLAIFLVTMVLGCELMSAVIMCATIAMLVNMFVGMWLMG 1146
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1161 ISYNAVSLINLVSAGMSVEFVSHYSTRSFAISKPTWLERAKEATISMGSAVAGVAMTN 1220
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSKGSRVERAEALAHMGSSVSGITLTK 1206
Qy      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1221 LFGILVLGLAKAQLTIQIFFRNLMLITLLGLHLGLVFLPVILSYVGPDVNPA 1272
Db      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1207 FGGIVVLAFAKSQIQIFVFRMYLAMVLLGATHGLIFLPVLLSYIGPSVNKA 1258

RESULT 59
US-10-170-205E-35088
; Sequence 35088, Application US/10170205E
; GENERAL INFORMATION:
; APPLICANT: ADAMS, Mark
; TITLE OF INVENTION: DEVICES, SUCH AS ARRAYS, COMPRISED OF HUMAN PROTEINS OR PROTEIN
; FILE OF INVENTION: CAPTURE AGENTS, AND USES THEREOF
; FILE REFERENCE: CLO01381
; CURRENT APPLICATION NUMBER: US/10/170,205E
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 40312
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 35088
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-170-205E-35088

Query Match      34.8%; Score 2402.5; DB 31; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy      7  RGWLWALLLRLLAQSEPYTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
Db      4  RGLALGLLILLLCPAQVFSQ-----SCVMYGECCIAYGDKRYNCEYSG----- 46
Qy      60 NTPARKITGDHLILLQKICPRLYTGNTQACCSAKOLVLEASISITKALLTRCPACSDN 119
Db      47 --PPKPLPKDGYDLVQELCPGFFG--NVSLCCDVQQLQTLKDNLQLPLQLFSLRCPFCFYN 103
Qy      120 FYNLHCHNTCPNQSLIFINVTR----VAQLGAGQLPAVVAEYAFYQHSFAEQSYDSCSRV 175

```

QY 1221 LPGILVILGAKAQLQIIFRNLNLTILGLHLGLVFLPVILSYGPDVNP 1272
DB 1207 FGIIVVLAFAKSIQIFIFRMYLAMVLLGATHGLIFLVLISYIGPSV 1258

RESULT 60
US-10-208-731-2
; Sequence 2, Application US/10208731
; GENERAL INFORMATION:
; APPLICANT: Carstean, et al.
; TITLE OF INVENTION: Genes for Niemann-Pick Type C Disease
; FILE REFERENCE: 4239-53894
; CURRENT APPLICATION NUMBER: US/10/208,731
; PRIOR FILING DATE: 2002-07-29
; PRIOR APPLICATION NUMBER: US/09/462,136
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: PCT/US98/13862
; PRIOR FILING DATE: 1998-07-02
; PRIOR APPLICATION NUMBER: US 60/051,682
; PRIOR FILING DATE: 1997-07-03
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-208-731-2

Query Match 34.8%; Score 2402.5; DB 32; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatch 450; Indels 103; Gaps 25;

QY 7 RGLLWALLRLAQSEPTTIIHQGYCAPYDECG-----KNPELSSGLMTLSNVCL 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCWYGECCGAYGDKRNCYSG-----46

QY 60 NTPARKITGDHLLILQKICPRLYTGNTQACSAKQLVSLEASLITKALTRCPACSDN 119
DB 47 --PPKPLPKGDYDLVQELCPGFFG-NVSLCCDVRLQTLKDLQLPLQLFSLRCPSCFN 103

QY 120 FVNLCHTNCSPNSQLFINVTR-----VAQLGAGQIPLAVVYEAFTQHGFAESYDSCSRV 175
DB 104 LNLFLCELTCSPROSQFLNVTATEDYDPTVNTQKTNKELQYVYVQGFANAMYNACRDV 163

QY 176 RVPAAATLAVGTCGYGSCALCNAQRNLNPFQDGTGNGLAPLDIT-----PHLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPPTITPVFSDFPVH-----215

QY 229 VGSIGIQLNEGVARNCSQDDVATCSQDCQCAASCPATARPQ-----ALDSTFYLG 279
DB 216 ---GMEPMNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDANYIM 272

QY 280 QMPSGLVLIILCSFVAVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVFGAFVAVCYKRYFSEYTPIDGNIASFV-NASDKGE-----ASCC 324

QY 327 DKLSFSTHLLQFQGGTGWASWPLTILVLSVTPVVALAAGLVFTLTTPDVLNAP 386
DB 325 DPVSAAFEGCLRLFRWGSCFCVRNPGCVIFPSLVFITACSSGLVFRVVTNPNVLDWAP 384

QY 387 NSQARSEKAFDQHPGPFRTNOVLTATPARNSSRYSDLSLLQPK-NFSGILDLDLLALL 445
DB 385 SSQARLEKEYFDQHPGPFRTNOVLTATPARNSSRYSDLSLLQPK-NFSGILDLDLLALL 444

QY 446 ELQERLRHLQWSPQAQRNLSIQDICYAPLNPDPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTI-ODICLAPLSYNI---NCITLSVLNLFQNSHSLVLDHKK 499

QY 506 NOTLMGQTSQVDWKDHFYLCANAPITFKDGTALALSCWADYGAVPFPFLATGGYKGDYS 565
DB 500 GDDFF---VIADYHTHLYCYRAPASLNDTSLLDHPCILGTGCGPFPVPLVGLGDDQNY 556

QY 566 EBAALIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMPQVTTFAERSLEDEI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQAWKEFINFVKNYKN---PNLTISFTAERSTED 613

QY 626 NETTAEDLPFATSYIVFLYISLALGSSYSSSRVWVDSKATILGLGGVAVLGAAMWG 685
DB 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDSKVSGLIAGILVILSSVACSLG 673

QY 686 PFSYLGRSSVLQVVPFLVSVGADNIFIFVLEYQRLPRPGBPREVHIGALGRVAP 745
DB 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQORDERLQGETLDQOLGRVLEVP 733

QY 746 SMLCSLSEAICFFLGALTPMVAFTALTGLAVILDLLQWAFVALLSDSKQREAS 805
DB 734 SMFLSSFSETVAFFLGALSVPVAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKQEK 793

QY 806 RLDVCCCKPQELPPGQ-GEGLLLGFPQKAYAPLHWHITRGVLLLLFLALFGVLSLYM 864
DB 794 RLDIFCCVRGABDGTSVQASESCLFRFKNSYSPILLKDDMRPIVIAIFVGLSPSIAVL 853

QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVYFVTTLYGNFSSSEAGNAICSSAG 924
DB 854 NKVDIGLDQSLMPDDSYAVDYFKISQVYLHAGPPVYFVLEEGHYTSSKGQNMVCGMG 913

QY 925 CNMFSFTQKIQYATEPPEQSYLAIPASSWVDDFIDWLT-SSCCRLYISGPNKDKFCPST 983
DB 914 CNMDSLVOQIPNAQLDNYTRIGFAPSSWIDDYFVWKPQSSCCRV---DNITDQFCNAS 970

QY 984 VNSLNLKNCMSIT-MGSVRPSVEQPHKVLPMFLNDRPNKCPKGLAAYSISVN--LTS 1040
DB 971 VVDPACVR-CRPLTPEGKQRPQGGDMRPLPMFLSDNPNKCPKGGHAAYSYSAVNILGH 1029

QY 1041 DQOVLSRPMVAHKPKNSQDYTEALRAARELAANITADLRKVPDPAFEVFPYTTITV 1100
DB 1030 GTRVATYFMYHTVLTQSADPIDALKKARLIASNT-ETMGINGS--AYRFPYSVFTV 1086

QY 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLNLLSIVMLIVDTVGFMAWD 1160
DB 1087 FVEQYLTIIDTIFNLVSLGAIPLVTVMLGCELWSAIVMCAITAMLVNMGVWMLWG 1146

QY 1161 ISYNAVSLINLVSAGMSVEFVSHITRSPALSTKPTWLERAKEATISMGSAVPAGVAMTN 1220
DB 1147 ISLNAVSLVNLVMSGIVSEFCFCSHITRAFTVSMKGSVERAEALAHMGSSVFSGITLTK 1206

QY 1221 LPGILVILGAKAQLQIIFRNLNLTILGLHLGLVFLPVILSYGPDVNP 1272
DB 1207 FGIIVVLAFAKSIQIFIFRMYLAMVLLGATHGLIFLVLISYIGPSV 1258

RESULT 61
US-10-218-140-5494
; Sequence 5494, Application US/10218140
; GENERAL INFORMATION:
; APPLICANT: Leach, Martin D.
; APPLICANT: Shimkets, Richard A.
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES
; TITLE OF INVENTION: ENCODED THEREBY
; FILE REFERENCE: 15966-543 CON
; CURRENT APPLICATION NUMBER: US/10/218,140
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 09/540,763
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/127,728
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/127,636
; PRIOR FILING DATE: 1999-04-02
; PRIOR APPLICATION NUMBER: 60/127,607
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 6322
; SOFTWARE: Curator Version 1.0
; SEQ ID NO 5494
; LENGTH: 1278
; TYPE: PRT

```

; ORGANISM: Homo sapiens
US-10-218-140-5494

Query Match      34.8%; Score 2402.5; DB 32; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLWALLLLRLAQSEPTTHHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCGAYGDKRYNCBSYG----- 46
QY 60 NTPARKITGDHLLLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG--NVSLCCDVRLQTLKDNQLPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSPNQSLFINVTR-----VAQLGAGQLPAPVAVAEFYQHSPAESYDSCSRV 175
DB 104 LLNLPCELTCSPROSQFNLVNTATEDYVDPVNTQTKNVKELQYYVVGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCGVYGSALCNAQRWLNFGQDTCGNGLAPLDT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA--CNATNWIEYMFNKDNGQAPFTTTPVFSDFPVH----- 215
QY 229 VGSQIOPNEGVARCNESQDDVATCSQDCAASCAPATARPQ-----ALDSTFYLG 279
DB 216 ---GMEPMNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILLCVSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLLVFGAFVAVCMYKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
QY 327 DXLSSTHTLLQFQCGHGTWASWPLTILVLSVTPVVALAAGLVTELTTPDELWSAP 386
DB 325 DPVSAAFEGCLRLTRMGSCFCVRNPGCVIFFSLVFITACSSGLVFVRTVTFNVDLWSAP 384
QY 387 NSQARSEKAFDQHPGPFRTNQVILTAPNRSRYSDSLGPK--NFGSILDLDLLELL 445
DB 385 SQAARLEKEYPDQHPGPFRTQVILTAPNRSRYSDSLGPK--NFGSILDLDLLELL 444
QY 446 ELQERLRLHQLVMSPEAQRNISIQDICYAPLNPNTSLVDCCTNSLLQYFQNNRTLLTLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSYNT---NCTILSVLNTFQNSHSLDHKK 499
QY 506 NQTLMGQTSQVDWKDHFELCANAPITFKDGTALALSCHADYCAPVPELAIGYKGDYS 565
DB 500 GDDFF---VYADYHFLYCVRAPASLNDTSLLDHPCLTGFGPVPVPLVGGYDDQNYN 556
QY 566 EBAELIMTSLNYPAGDPRLAQAKLWEAPLEENRAFORMMAGMFQVTFTAERSLEDBI 625
DB 557 NATALVITFPVNNYNDTEKLRQAQWEKEFINFKVKNYK---PNLTISFTAERSIEDEL 613
QY 626 NRTTADBLPFIATSYIVIFLYTSLALGVSYSMSRWVDSKATLGGVAVVLGAVMAAMG 685
DB 614 NRESDDVTWVISYAIMFLYISLALGHKSCRRLLVDSKVSIGTAGILVILSVVACSIG 673
QY 686 FSYLGIRSIIVLQVPELVLSVGDNITFI VLEYQRLPRPGEPRVHIGRALGRVAP 745
DB 674 VFSYIGLPLUTLIVIEIPFLVAVGVNDNIFILVQAYORDERLQGETLQDLQGRVJGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMAVRTFALTSGLAVIDLFLQWSAFVALLSDSKRQES 805
DB 734 SMFLSFSFTVAFGLSVMRAVHTFSLFAGLAVFIDELLQITCFVSLGLDIXRQEK 793
QY 806 RLDDCCVKQBLPPPGQ--GEGLLGFFQKAVAPFLHWHITRGVULLFLALFGVLSLYM 864
DB 794 RLDFICVGAEDGTSVQASESCLFRFFKNYSPLLLKDWMRPIVIAIFVGLSFSIAVL 853
QY 865 CHISVLQDELAPKDSYLLDFLNRIFYEGCAPVYFVTLTGYNFSSAGNACSSAG 924
DB 854 NKVDIGLQDSLMPDSDYVWDYFKSISQYLHAGPPVYVLEEGHDYTSKGQNMVCGMG 913
QY 925 CNNFSTQKIQVATEPPEQSILAIPASSWVDIDWLTP--SSCCRLIYISGNPKDFCPST 983
DB 925 CNNFSTQKIQVATEPPEQSILAIPASSWVDIDWLTP--SSCCRLIYISGNPKDFCPST 983

RESULT 62
US-10-741-600-1542
; Sequence 1542, Application US/10741600
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1542
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-600-1542

Query Match      34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLWALLLLRLAQSEPTTHHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECCGAYGDKRYNCBSYG----- 46
QY 60 NTPARKITGDHLLLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG--NVSLCCDVRLQTLKDNQLPLQFLSRCPSCFYN 103
QY 120 FVNHLCHNTCSPNQSLFINVTR-----VAQLGAGQLPAPVAVAEFYQHSPAESYDSCSRV 175
DB 104 LLNLPCELTCSPROSQFNLVNTATEDYVDPVNTQTKNVKELQYYVVGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCGVYGSALCNAQRWLNFGQDTCGNGLAPLDT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA--CNATNWIEYMFNKDNGQAPFTTTPVFSDFPVH----- 215
QY 229 VGSQIOPNEGVARCNESQDDVATCSQDCAASCAPATARPQ-----ALDSTFYLG 279
DB 216 ---GMEPMNNATKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYIM 272
QY 280 QMPGSLVLIILLCVSFVAVVTILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLLVFGAFVAVCMYKRYFVSEYTPIDSNIAFSV--NASDKGE-----ASCC 324
QY 327 DXLSSTHTLLQFQCGHGTWASWPLTILVLSVTPVVALAAGLVTELTTPDELWSAP 386
DB 325 DPVSAAFEGCLRLTRMGSCFCVRNPGCVIFFSLVFITACSSGLVFVRTVTFNVDLWSAP 384
QY 387 NSQARSEKAFDQHPGPFRTNQVILTAPNRSRYSDSLGPK--NFGSILDLDLLELL 445
DB 385 SQAARLEKEYPDQHPGPFRTQVILTAPNRSRYSDSLGPK--NFGSILDLDLLELL 444
QY 446 ELQERLRLHQLVMSPEAQRNISIQDICYAPLNPNTSLVDCCTNSLLQYFQNNRTLLTLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDCLAPLSYNT---NCTILSVLNTFQNSHSLDHKK 499
QY 506 NQTLMGQTSQVDWKDHFELCANAPITFKDGTALALSCHADYCAPVPELAIGYKGDYS 565
DB 500 GDDFF---VYADYHFLYCVRAPASLNDTSLLDHPCLTGFGPVPVPLVGGYDDQNYN 556
QY 566 EBAELIMTSLNYPAGDPRLAQAKLWEAPLEENRAFORMMAGMFQVTFTAERSLEDBI 625
DB 557 NATALVITFPVNNYNDTEKLRQAQWEKEFINFKVKNYK---PNLTISFTAERSIEDEL 613
QY 626 NRTTADBLPFIATSYIVIFLYTSLALGVSYSMSRWVDSKATLGGVAVVLGAVMAAMG 685
DB 614 NRESDDVTWVISYAIMFLYISLALGHKSCRRLLVDSKVSIGTAGILVILSVVACSIG 673
QY 686 FSYLGIRSIIVLQVPELVLSVGDNITFI VLEYQRLPRPGEPRVHIGRALGRVAP 745
DB 674 VFSYIGLPLUTLIVIEIPFLVAVGVNDNIFILVQAYORDERLQGETLQDLQGRVJGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMAVRTFALTSGLAVIDLFLQWSAFVALLSDSKRQES 805
DB 734 SMFLSFSFTVAFGLSVMRAVHTFSLFAGLAVFIDELLQITCFVSLGLDIXRQEK 793
QY 806 RLDDCCVKQBLPPPGQ--GEGLLGFFQKAVAPFLHWHITRGVULLFLALFGVLSLYM 864
DB 794 RLDFICVGAEDGTSVQASESCLFRFFKNYSPLLLKDWMRPIVIAIFVGLSFSIAVL 853
QY 865 CHISVLQDELAPKDSYLLDFLNRIFYEGCAPVYFVTLTGYNFSSAGNACSSAG 924
DB 854 NKVDIGLQDSLMPDSDYVWDYFKSISQYLHAGPPVYVLEEGHDYTSKGQNMVCGMG 913
QY 925 CNNFSTQKIQVATEPPEQSILAIPASSWVDIDWLTP--SSCCRLIYISGNPKDFCPST 983
DB 925 CNNFSTQKIQVATEPPEQSILAIPASSWVDIDWLTP--SSCCRLIYISGNPKDFCPST 983
```

```
QY 387 NSQARSEKAFHDQHPGPFPTNQVILTAAPNRSSRYRDSLLLGPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKBYFDHQHGFPTTEQLIIRAPLTDKHIYQVPBGADVPFGPLDIQILHQLV 444
QY 446 ELQERLRHLQVMSPEAQRNLSLODICVAPLNPNTSLYDCINSLQYFQNNETLALLTA 505
Db 445 DLQIAIEN--ITASYNDETWTQIDICLAPLSPYNT--NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKDGOTALALSCMADYDAPVPFPFLAIGYKGDYS 565
Db 500 GDDPFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFGGVPFPWLVLGGYDQNYN 556
QY 566 EAEALMTFSLNYPAGDPRLAQAKLWEBAFLWEAFQRRMAGMFQVTPAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLRQAQAWKEFNFVKNYKN--PNLTISPTAERSLEDEL 613
QY 626 NRTTAEDELPIFATSYIVFIYLSALGSYSSMSRVMVDSKATGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRLLVDSKVSIGIAGILVLSSVACSLG 673
QY 686 FFSYLGRSLVILQVVPFLVAVSGADNIPFVLEYQRLPRRPGEPREVIHIGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFLVLAVGVNDIFILVQAYQORDERLQGETLDQQLGRVLEAVP 733
QY 746 SMLLCSLSAICFFLGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSSFSSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLDQKQKN 793
QY 806 RLDVCCVQKQELPPGQ--GEGILLGFQKAYAPFLHWTIRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNYSVPLLLKDMRPIVIAIFVGLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVFTVTLGYNFSSBAGMNAICSSAG 924
Db 854 NKVDIGLDQSLNPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDTSSKGQNMVCGMG 913
QY 925 CNNFSTQKIYATEPEPSOYLAIAPASSWDDFIDLWLP--SSCCRLYISGPNKDKFCPST 983
Db 914 CNNDLSVQOIFNAALDNTYTRIGFAPSSWIDDYFDMVKPQSSCCRV--DNITDQFCNAS 970
QY 984 VNSLNCCLKMCSIT--MGSVRPSVEQHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPAQVR--CRPLTPGKQRPQGGDFMFLPMLSDPNPNPKCGKGAAYSSAVNTLLGH 1029
QY 1041 DGQVLASRFMAWHKPKNSQDYTEALRAARELAANITADLRKVPYGTDPRAFEVPPYTTINV 1100
Db 1030 GTRVGATYFMTYHTVLOTGADFIDALKARLIASNT--ETMGINGS--ATRVFPYSVTV 1086
QY 1101 FVEQYLTIPEGLFMLSCLVPTFAVSCLLLGLDLSGLNLLSIVMILVDTVGFMAWMD 1160
Db 1087 FVEQYLTIIDDTIFNLGVSIGALFVMTVLGCELWSAVIMCATIAMVLNMGVNMWLMG 1146
QY 1161 ISYNAVSLNLVSAVMSVEFVSHITRSFAISTYKPTWLERAKATISMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVSMKGSRVERABEALAHMGSSVFSGITLTK 1206
QY 1221 LRGILVLGAKAQLOIOPFRLNLLATLGLHGLVPLPVLSYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSIQIFIFRMYLAMVLLGATHGLIPLVLLSYIGPSVNKA 1258
```

RESULT 63

```
US-10-741-601-530
; Sequence 530, Application US/10741601
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 530
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-530
```

```
Query Match 34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGWLLWALLRLAQSEPYTHIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLALGLLILLCLCPAFYFQ-----SCWYGEIGIAYGKRYNCEYS----- 46
QY 60 NTPARKITGDHILLLKICPRLYTGPTQACSAKOLVSLKASLSTTKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVDELCEGFFG--NVSLLCCDVRLQTLKONLQPLQFLSKCPSCFYN 103
QY 120 FVNLHCHNTCSNQSLFIVNTR----VAQIGAGQLPAVWAYEAFYQHSFAESYDSCSRV 175
Db 104 LLNLFCELTCSQRSQFLNVTATEDYVDPTNQTNTKTVKELQYVVGQSPANAMYACRDV 163
QY 176 RYPAATAVAVTMCVGYGSALCNAQRWLAFOGDTGNGLAPLDIT-----FHLBPGQA 228
Db 164 EAPSSNDKALGLLGCXKADA--CNATWIEYMFKNKGQAPFTTITPVFSDPPVH----- 215
QY 229 VSGGIQPLNEGVARCNESQGDVATCSCODCAASCAPAIARPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATYKGCDSDEVTAPCSQDCSIVCGPKQPPPPPPAPWTILGDAMVIM 272
QY 280 QMPGSLVLIILCSVFVAVVTILL-----VGRVAPAPARKSKMVDPKGTSL 326
Db 273 WITYMAFLVFGAPFAVVCYKRYFVSEVTPIDSNIAFSV--NASDKGE-----ASCC 324
QY 327 DKLSPTHTYLQOPFGCGTWTWASWPLTILVLVSVIPVVALAGLVTELTDDPVELMSAP 386
Db 325 DFVSAFEGCLRLRLETRWGSFCVRNPGCVIFPSLVEITACSSGLVFRVTVNTEVDLMSAP 384
QY 387 NSQARSEKAFHDQHPGPFPTNQVILTAAPNRSSRYRDSLLLGPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKBYFDHQHGFPTTEQLIIRAPLTDKHIYQVPBGADVPFGPLDIQILHQLV 444
QY 446 ELQERLRHLQVMSPEAQRNLSLODICVAPLNPNTSLYDCINSLQYFQNNETLALLTA 505
Db 445 DLQIAIEN--ITASYNDETWTQIDICLAPLSPYNT--NCTILSVLNYFQNSHVLDDHK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKDGOTALALSCMADYDAPVPFPFLAIGYKGDYS 565
Db 500 GDDPFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFGGVPFPWLVLGGYDQNYN 556
QY 566 EAEALMTFSLNYPAGDPRLAQAKLWEBAFLWEBAFQRRMAGMFQVTPAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLRQAQAWKEFNFVKNYKN--PNLTISPTAERSLEDEL 613
QY 626 NRTTAEDELPIFATSYIVFIYLSALGSYSSMSRVMVDSKATGLGGVAVVLGAVMAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRLLVDSKVSIGIAGILVLSSVACSLG 673
QY 686 FFSYLGRSLVILQVVPFLVAVSGADNIPFVLEYQRLPRRPGEPREVIHIGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFLVLAVGVNDIFILVQAYQORDERLQGETLDQQLGRVLEAVP 733
QY 746 SMLLCSLSAICFFLGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKRQEAS 805
Db 734 SMFLSSFSSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLQITCFVSLGLDQKQKN 793
QY 806 RLDVCCVQKQELPPGQ--GEGILLGFQKAYAPFLHWTIRGVLLFLALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNYSVPLLLKDMRPIVIAIFVGLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFEVGAPEVFTVTLGYNFSSBAGMNAICSSAG 924
Db 854 NKVDIGLDQSLNPDSDSYMDYFKSISQYLHAGPPVYFVLEEGHDTSSKGQNMVCGMG 913
QY 925 CNNFSTQKIYATEPEPSOYLAIAPASSWDDFIDLWLP--SSCCRLYISGPNKDKFCPST 983
Db 914 CNNDLSVQOIFNAALDNTYTRIGFAPSSWIDDYFDMVKPQSSCCRV--DNITDQFCNAS 970
QY 984 VNSLNCCLKMCSIT--MGSVRPSVEQHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPAQVR--CRPLTPGKQRPQGGDFMFLPMLSDPNPNPKCGKGAAYSSAVNTLLGH 1029
QY 1041 DGQVLASRFMAWHKPKNSQDYTEALRAARELAANITADLRKVPYGTDPRAFEVPPYTTINV 1100
Db 1030 GTRVGATYFMTYHTVLOTGADFIDALKARLIASNT--ETMGINGS--ATRVFPYSVTV 1086
QY 1101 FVEQYLTIPEGLFMLSCLVPTFAVSCLLLGLDLSGLNLLSIVMILVDTVGFMAWMD 1160
Db 1087 FVEQYLTIIDDTIFNLGVSIGALFVMTVLGCELWSAVIMCATIAMVLNMGVNMWLMG 1146
QY 1161 ISYNAVSLNLVSAVMSVEFVSHITRSFAISTYKPTWLERAKATISMGSAVPAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVSMKGSRVERABEALAHMGSSVFSGITLTK 1206
QY 1221 LRGILVLGAKAQLOIOPFRLNLLATLGLHGLVPLPVLSYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSIQIFIFRMYLAMVLLGATHGLIPLVLLSYIGPSVNKA 1258
```

QY 925 CNNEFTOKIQVATEPPEQSYLAIPASSWDDDFIDWLTP--SSCRLYISGPNKQKPCPST 983
DB 914 CNDSLVQOIFNAAQLDNVTIRIGFAPSSWIDDYFDWVRPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT--MGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VDPACVR--CRPLTPGKORPQGGDMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQVLAASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
DB 1030 GTRVGATYFMYHTVTLQTSADFDALKKARLIASNT--ETMINGS--AYRVFPYSVFV 1086
QY 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCILGLDLRSGLNLLSVMLIVDTVGPMALWD 1160
DB 1087 FVEQYLTITIDDTIFNLGSLGAIFLVWVLLGCELSAVIMCATIAMVLNMGFWMLWG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAISKTWLERAKEATISMGSAVAGVAMTN 1220
DB 1147 ISLNAVSLNLMVSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
QY 1221 LFGIIVLGLAKAQLQIIFFRNLNLLITLGLLHGLVFLPVILSYVGPDPNPA 1272
DB 1207 FGGIIVLAFAKSIQIFVFRMYLAMVLLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 64
US-10-756-149-4924
; Sequence 4924, Application US/10756149
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: NOVEL METHODS OF DIAGNOSIS OF METASTATIC CANCER, COMPOSITIONS AND
; TITLE OF INVENTION: METHODS OF SCREENING FOR MODULATORS OF METASTATIC CANCER
; FILE REFERENCE: file
; CURRENT APPLICATION NUMBER: US/10/756,149
; NUMBER OF SEQ ID NOS: 5818
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4924
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-756-149-4924

Query Match 34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLAQSEPTTIHQPGVCAPYDECG-----KNPELSSGLMTLSNVCLLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCWYGGCGIAYGDKRYNCEVSG----- 46
QY 60 NTPARKITGDHLILQKICPRLYTGTQNTQACCSAKQLVSLASLSITKALLRCPACSDN 119
DB 47 --PPRLPKDGYDLVQELCPGFFG--NVSLCCDVRLQTLKDNILQLPLQFLSCRCFCFN 103
QY 120 FVNLHCHNTCSNQSLFINVTR-----VAQLGAGQIPAVVAYEAFYQHSFAESYDSCSRV 175
DB 104 LNLFLCELTCSPRSQSLNVTATEDYVDPVTNQTKNVKELQYVYVQGSFANAMYNACRDV 163
QY 176 RYPAATLAVGTCMGVYGSALNAQRLNFGQDTGNGLAPLDT-----PHLEPGQA 228
DB 164 EAPSNNDKALGLLCKGADA--CNATNWIEMFNKONGQAPFTITPVFSDFPVH----- 215
QY 229 VGSGLQPLNEGVARCNESQDDVATCSQDCAASCPAIAAPQ-----ALDSTFYLG 279
DB 216 ---GHEPMNNAKGCDESVDVATPCSCDCSIVCGPKQPPPPAPWILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFPAVVTILL-----VGRFVAPARDKSKMDPKGKTSLS 326
DB 273 WITYMAFLLVFGAPFAVVCYRKRYFVSBYTPTIDSNIAFSV--NASDKGE-----ASCC 324

QY 327 DKLSFSTHTLLGQFPQGGTWTWASWPLTILVLSVIPVVALAAGLVFTLTDTDPVLSWAP 386
DB 325 DPVSAAPFEGCLRLRLLFTRWGSFCVRNPGCVIFPSLVFITACSSGLVFRVTINPVDLWSAP 384
QY 387 NSQARSEKAFHQHFGPPFRFNQVILTAPNRSSVRYDSLLILGPK--NFGSILDLDLLELL 445
DB 385 SSOARLEKEYDQHFPGPFRTPEQLIIRAPLDKHIYQYPGADVPFGPPLDIQLHQLV 444
QY 446 ELQERLRLHQLWSPEAQNRNISLQDICVAPLNDPNTSLYDCCINSILQYFQNNRTLLLLTA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDILCLAPLSPYNT---NCTILSVLNTFQNSHVDHKK 499
QY 506 NQTLMGQTSQVDWKDHFLYCANAPLTPFKDGTALALSMDADYCAPFPPFLAIGYKGKDY 565
DB 500 GDDPF--VYADYTHFLYCYRAPASLNDTSLHDPCLGTFGPGVFPMLVGGYDQDYN 556
QY 566 BAEALIMFTSLNTPAGDPRLAQAKLWBEAFLEENRAFQORMAGMFQVTFTAERSLEDBI 625
DB 557 NATALVITFPVNNYNDTEKLQRAQWKEKTFNFVKNYKN---PNLTISFTAERSTEDL 613
QY 626 NRTTAEDLPFATSYIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAYMAAG 685
DB 614 NRESQSDVTVVTSYAIMFLYISLALGHIKSCRLLVDSKVSGLAGILIVLUSSVACSUG 673
QY 686 FFSYLGIIRSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHIGRALGRVAP 745
DB 674 VFSYIGLPLTIVIEVIFFLVLANGVNDIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
QY 746 SMLLCSLSEAJCPFLGALTMPAVNTPALTGLAVILDPLQMSAFVALLSLSKROEAS 805
DB 734 SMFLSSFSETVAFFLGALSVMFAVHTFSLFAGLAVFIDFLQLITCFVSLGLDIKRQEN 793
QY 806 RLDVCCVCPQELPPPGQ--GEGLLGFFQKAYAPELLHWITRGVLLPLFALFGVLSYM 864
DB 794 RLDIFCCVCRGAEADGTSVQASECLFRFPKNSVSPLLKDMWRPIVIALFVGLSVLSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLPLNRYFEVGPVYFVTLGYNFSSAGNAICSSAG 924
DB 854 NKVDIGLQSLSMPPDDSYMDYFKSISOYLHAGPPVYFVLEBHDYTSKQGNMVGCGMG 913
QY 925 CNNEFTOKIQVATEPPEQSYLAIPASSWDDDFIDWLTP--SSCRLYISGPNKQKPCPST 983
DB 914 CNDSLVQOIFNAAQLDNVTIRIGFAPSSWIDDYFDWVRPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT--MGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
DB 971 VDPACVR--CRPLTPGKORPQGGDMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQVLAASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVPPYITNV 1100
DB 1030 GTRVGATYFMYHTVTLQTSADFDALKKARLIASNT--ETMINGS--AYRVFPYSVFV 1086
QY 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCILGLDLRSGLNLLSVMLIVDTVGPMALWD 1160
DB 1087 FVEQYLTITIDDTIFNLGSLGAIFLVWVLLGCELSAVIMCATIAMVLNMGFWMLWG 1146
QY 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAISKTWLERAKEATISMGSAVAGVAMTN 1220
DB 1147 ISLNAVSLNLMVSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
QY 1221 LFGIIVLGLAKAQLQIIFFRNLNLLITLGLLHGLVFLPVILSYVGPDPNPA 1272
DB 1207 FGGIIVLAFAKSIQIFVFRMYLAMVLLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 65

US-10-796-280-1261

; Sequence 1261, Application US/10796280

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001510


```
; CURRENT APPLICATION NUMBER: US/10/796,280
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 68533
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1261
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-796-280-1261

Query Match          34.8%; Score 2402.5; DB 37; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGMWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEVSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKDNQLPLQLFLSRCPSCFYN 103

QY 120 FVNLCNTCSNQSFLINVTNTR-----VAQLGAGQLPAAVVAEAFYQHSFABQSYDSCSRV 175
DB 104 LNLFLCBLTCSRQSQFLNVTATEDYDVPVNTQTKNVKELQYVVGQSPANAMYNACRDV 163

QY 176 RYPAATLAVGTMCGYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEYMFNKDNGQAPPTITPVFSDFPVH-----215

QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPNMNAKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYVIM 272

QY 280 QMPSGLVLIILLCSPFVVVITILL-----VGRVAPARDKSKMVDPKKGTSL 326
DB 273 WITYMAFLVFGAFVAVWCYKRYFVSEYTPIDSNIAFSV-NASDKG-----ASCC 324

QY 327 DKLSFSTHTLLQFGQMGTVASWPLTTLVLSVPPVVALAAGLVFTLTTPDVELWSAP 386
DB 325 DPVSAAPGECRLRLTRWGSFCVRPGCVIFSLVPIFACSSGLVFRVTNTPVLDWSAP 384

QY 387 NSQARSEKAFDHPGPPRTNQVILTAPNRSRYDLSLLGPK-NFSGILDLDLLELL 445
DB 385 SSOARLEKEFYDHPGPPRTQELIIRAPLTKHIYQYPVPGADVPFGPLDIQILHVL 444

QY 446 ELQERLRHLQWSPENARQNISLQDIYAPLNPDNTSLYDCCINSLLQYFQNNRTLLITA 505
DB 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVLNYFQNSHVLHKK 499

QY 506 NQTLMGQTSQVDKDHFLYCANAPLTFKDGTAALSCMADYGAAPVFPFLAIGYKGDYS 565
DB 500 GDDFPF---VYADYHFLYCVAPASLNDTSLHDPCLGTGFGPFPFLVGLGYDDQYN 556

QY 566 EBAALIMTFSLNYPAGDPRLAQAKLEBAPLEEMRAFORRMAGMFQVTTFAERSLEDBI 625
DB 557 NATALVITPVNNYNDTEKLQRAQAKEFNFVKYKN---PNLTISFTAERSIEDBL 613

QY 626 NRTTAEPLPIFATSVIVIFLYLSLALGYSYSSRWVSDSKATLGGVAVVLGAVNAGM 685
DB 614 NRESDSVFVTVISVAIMEFLYLSLALGHITKSCRRLLVDSKVSIGLAGILVLSVACSLG 673

QY 686 PFSYLGRSLVILQVPPVLVLSVGADNIFI VLEYQRLPRPRGPREVHIGRALGRVAP 745
DB 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQDERLQGETLQDLQGLSERVGEVAP 733

QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKRQEAS 805
DB 734 SMFLSSFSFETVAFGLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIDKQEK 793

QY 806 RLVDCCVCKPQBLPPPGQ--GEGILLGFFQKAYAPFLHWTIRGVVLLFLALFGVLSY 864
DB 794 RLIDFCCVRGABDGTSVQASECLFRFFKNYSPLLLKDMWRPIVIAIFVGLVLSFI 853
```

RESULT 66

```
US-10-796-307-810
; Sequence 810, Application US/10796307
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001509
; CURRENT APPLICATION NUMBER: US/10/796,307
; CURRENT FILING DATE: 2004-03-10
; NUMBER OF SEQ ID NOS: 44201
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 810
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-796-307-810
```

Query Match 34.8%; Score 2402.5; DB 37; Length 1278;

Best Local Similarity 39.8%; Pred. No. 2.3e-201;

Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

```
QY 7 RGMWALLRLRLAQSEPTTHIQPGYCAFYDECG-----KNPELSGLMTLSNVCSLS 59
DB 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEVSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVLEASLSITKALLTRCPACSDN 119
DB 47 --PPKPLPKDGYDLVQELCPGFFG-NVSLCCDVRQLQTLKDNQLPLQLFLSRCPSCFYN 103

QY 120 FVNLCNTCSNQSFLINVTNTR-----VAQLGAGQLPAAVVAEAFYQHSFABQSYDSCSRV 175
DB 104 LNLFLCBLTCSRQSQFLNVTATEDYDVPVNTQTKNVKELQYVVGQSPANAMYNACRDV 163

QY 176 RYPAATLAVGTMCGYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGQA 228
DB 164 EAPSSNDKALGLLCKGDADA-CNATNWIEYMFNKDNGQAPPTITPVFSDFPVH-----215

QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
DB 216 ---GMEPNMNAKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTILGLDAMYVIM 272

QY 280 QMPSGLVLIILLCSPFVVVITILL-----VGRVAPARDKSKMVDPKKGTSL 326
```



```

Db      614  NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDSKVSIGIAGILVLSVACSGLG 673
Qy      686  PFSYLGRSSVLVQVVPFLVSGADNIPFVLEYQRLPRRPGEBPREVHIGRALGRVAP 745
Db      674  VPSYIGLPLTLVIEVIFPLVAVGVNIFILVQAYQDERLQGETLDQQLGRVLEVP 733
Qy      746  SMLLCSLSEACFFIGALTPMPAVRTPALTSLGLAVLDLFLQMSAFVALLSLDSKQREAS 805
Db      734  SMLFSSFSFETVAFPLGALSMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIDKQREKN 793
Qy      806  RLDDVCCVKPQBLPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLFALFGVSLYSM 864
Db      794  RLDFCCVGAEDGTSVQASBSCLFRFFKNSYSPLLKDWMPRIVIAIFVGLSFSIAVL 853
Qy      865  CHISVGLDQELALPKDSYLLDYLFLNRYFEVGAPVYFVTLGYNFSSBAGMAICSSAG 924
Db      854  NKVDIGLQSLSMPPDSDYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKGGQNMVCGMG 913
Qy      925  CNPFSFTQKIQYATEPPEQSYLAIPASSWDDFIDWLTP-SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDLSVQOIFNAQAQLDNRYTRIGFAPSSWIDYDFWVKPQSSCCRV---DNITDQFCNAS 970
Qy      984  VNSLNLCKNCSIT-MGSVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db      971  VDDPACVR-CRPLTPEGKORPOGGDFMRFLPMFLSDNPNPKCKGKGHAAYSSAVNTLLGH 1029
Qy      1041  DGQVLASRFMAVHKPKNSQDYTEALRAARELAANITADLRKVPGTDPAPFEVPPYITNV 1100
Db      1030  GTRVGATYFTVHTVLOTSAFDALKKARLIASNT-ETMGINGS--AYRFPYSVYV 1086
Qy      1101  FFEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLSGLLNLISIVMILVDTVGFMAWD 1160
Db      1087  FFEQYLTLIDDTIFNLGVSIGALFLVTMVLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146
Qy      1161  ISNAVSLNLVSAGVMSVEFVSHITSPASTKPTWLERAKEATISMSAPGAVMTN 1220
Db      1147  ISLNAVSLNLVNSCGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLTK 1206
Qy      1221  LRGILVLGLAKAQIQLPFRLNLTLTLGLLHGLVFLPVLSYVGPDPNPA 1272
Db      1207  FGGIVVLAKSQIFQIFRYMYLAWVLGATHGLIFLVLVLSYIGPSVKA 1258

```

RESULT 68

US-10-995-561-952

; Sequence 952, Application US/10995561

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele et al.

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; TITLE OF INVENTION: DETECTION AND USES THEREOF

; FILE REFERENCE: CL001559

; CURRENT APPLICATION NUMBER: US/10/995,561

; CURRENT FILING DATE: 2004-11-24

; NUMBER OF SEQ ID NOS: 85702

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 952

; LENGTH: 1278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-995-561-952

```

Query Match      34.8%; Score 2402.5; DB 39; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2,3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

```

```

Qy      7  RGWLWALLRLAQSPYTTIHQPGYCAFYDEG-----KNPELSSGLMTLSNVCSLS 59
Db      4  RGALGLLALLLCPAQVFSQ-----SCWYEGGAYGDKRYNCEYISG-----46
Qy      60  NTPARKITGDHILILQKICPRLYTGNTQACCSAKQLVSEASLSITKALLTRCPACSDN 119

```

```

Db      47  --PPKPLPKGDYDLVQELCPGFFFG-NVSLCCDVRLOLTLDKDLQPLQLFSLRCPSCFYN 103
Qy      120  FVNHLCHNTCSNPQSLFINVTR----VAOLGAGQLPAVVAAYFAFYQHSFAEQYSDCSR 175
Db      104  LLNLFCELTCSPROSQFLNVATATEDYVDPYTNQTKTNKVELQYVQGSFANMYNACRDV 163
Qy      176  RVPAAATLAVGTWCGVYGSALCNAQRLNFPQDGTGNGLAPLDIT-----FHLLEPGQA 228
Db      164  EAPSSNDKALGLLCKGDADA-CNATNWIYMFKNKGQAPFTTITPVSPDPVH-----215
Qy      229  VSGIQLPNEGVARCNESOGDDVATCSODCAASCAPAIARPO-----ALDSFPLG 279
Db      216  ---GMEPMNATKGCDSDEVTAPCQDCSIVCGPKQPPPPPPAPWTILGLDAMVIM 272
Qy      280  QMPGSLVLIILCSVPVAVTILL-----VGRVAPARDKSMVDPKKGTSL 326
Db      273  WITYMAFLVFFGAPFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy      327  DKLSFSTHTLLQOFFPGWGTWVASWPLTILVLSVIVPVALAAGLVFTLTDTPVELWSAP 386
Db      325  DPVSAAFEGCLARLFTRMGSFCVRNPGCVIFTSLVFITACSSGLVFRVVTNPNVDLWSAP 384
Qy      387  NSQARSEKAPHQHPGPPRTNQVILTAPNRSYRYSLLGPK-NFSGILDLDLLELL 445
Db      385  SSOARLEKEYFDQHPGFFRTEQLIIRAPLTDKHIYQYPSPGADVPFGPPLDIQILHQVL 444
Qy      446  ELQERLRLQVMSPEAQRNLSQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db      445  DLQIAIEN--ITYASDNEVTIQLDCLAPLSPYNT---NCTILSVNLNYFQNSHVSIDHKK 499
Qy      506  NOTLMGQTSQVMDKOHFLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGYKGDYS 565
Db      500  GDDFF---VYADVHTFLYCVRAPASLNDTSLDHPCLGTFGGVPFVFWLGLGYDDQNYN 556
Qy      566  EBAELIMTSLNYPAGDPRLAQAQKWEAFLEEMRAPFORRMAGMQVFTFAERSLEDEI 625
Db      557  NATALVITPVVNYNDTEKLQRAQAWEKEFINFKVKNYK---PNLTIPTFAERSLEDEL 613
Qy      626  NRTTARDLPIFATSYIVIFLYISLALGSYSSNMVDSKATLGLGVAVLGAVMAAMG 685
Db      614  NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDSKVSIGIAGILVLSVACSGLG 673
Qy      686  PFSYLGRSSVLVQVVPFLVSGADNIPFVLEYQRLPRRPGEBPREVHIGRALGRVAP 745
Db      674  VPSYIGLPLTLVIEVIFPLVAVGVNIFILVQAYQDERLQGETLDQQLGRVLEVP 733
Qy      746  SMLLCSLSEACFFIGALTPMPAVRTPALTSLGLAVLDLFLQMSAFVALLSLDSKQREAS 805
Db      734  SMLFSSFSFETVAFPLGALSMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIDKQREKN 793
Qy      806  RLDDVCCVKPQBLPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLFALFGVSLYSM 864
Db      794  RLDFCCVGAEDGTSVQASBSCLFRFFKNSYSPLLKDWMPRIVIAIFVGLSFSIAVL 853
Qy      865  CHISVGLDQELALPKDSYLLDYLFLNRYFEVGAPVYFVTLGYNFSSBAGMAICSSAG 924
Db      854  NKVDIGLQSLSMPPDSDYMDYFKSISQYLHAGPPVYFVLEEGHDYTSKGGQNMVCGMG 913
Qy      925  CNPFSFTQKIQYATEPPEQSYLAIPASSWDDFIDWLTP-SSCCRLYISGPNKDKFCPST 983
Db      914  CNNDLSVQOIFNAQAQLDNRYTRIGFAPSSWIDYDFWVKPQSSCCRV---DNITDQFCNAS 970
Qy      984  VNSLNLCKNCSIT-MGSVRPSVEQFHKYLPMFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db      971  VDDPACVR-CRPLTPEGKORPOGGDFMRFLPMFLSDNPNPKCKGKGHAAYSSAVNTLLGH 1029
Qy      1041  DGQVLASRFMAVHKPKNSQDYTEALRAARELAANITADLRKVPGTDPAPFEVPPYITNV 1100
Db      1030  GTRVGATYFTVHTVLOTSAFDALKKARLIASNT-ETMGINGS--AYRFPYSVYV 1086
Qy      1101  FFEQYLTLPEGLFMLSCLVPTFAVSCLLGLDLSGLLNLISIVMILVDTVGFMAWD 1160
Db      1087  FFEQYLTLIDDTIFNLGVSIGALFLVTMVLGCELWSAVIMCATIAMVLVNMFGVMWLMG 1146

```


Db 500 GDDFF---VYADYHTFLYCVRAPASLNDTSLHDPCLGTGCGPVFPMLVLGGYDQNYN 556
Qy 566 EEAALIMTFSLNYPAGDPRLAQAKLWEEAFLEEMRAFORMAGMFOVTFTEARSLDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWKEKFNFVKNYKN---PNLTISFTAERSTIEDEL 613
Qy 626 NRTTAEDLPFATSVIVIFLYISLALGSYSSSRVMDSKATLGLGGVAVVLGAYMAAMG 685
Db 614 NRESDSVFTVVISVAINFELYSIALGHKSCERLLVDSKVSGLTAGILVLSSVACSIG 673
Qy 686 FFSYLGIIRSLVILQVPPFLVSVGADNIFIFVLEYQRLPRPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVLPFLVAVGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
Qy 746 SMLCSLSEAICFFLGALTPMPAVTFTALTSGLAVIDLFLQMSAFVALLSDSKROEAS 805
Db 734 SMLFSFSETVAFILGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQEN 793
Qy 806 RLDVCCVKPQELPPPGQ-GEGLLIGFFOKAVAPFLHWHITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSVSPLLDKDMPPIVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNPFSSAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDSDSYMDVYFKSISQYLHAGPPVYFVLEGRHDYTSKGGQNMVCGG 913
Qy 925 CNPFSTQKIQYATEFPEQSYLAIAPASSWDDFIDWLT-SSCCRLYISGPNKDECPST 983
Db 914 CNNDSLVQOIFNAALQDNTRYIGFAPSSWIDDYFDWVKPQSSCCRY---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT-MGSVRPSVEOFHKYLPFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLETPEGKORPOGGDFMRFLPFLSDNPNPKCKGGAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFPPYITNV 1100
Db 1030 GTRVGATYEMTYHTVLTQTSADFDALKARLIASNTV-ETMGINGS--AYRVEPYSVFY 1086
Qy 1101 FVEQYLTILPEGLFMLSCLIVTEFAVSCLLIGLDLSGLNLLSIVMLIVDTVGPALWD 1160
Db 1087 FVEQYLTIIIDDYTFNIGVSLGAIFLVTMVLGCELWSAIVMCATTAMVLNFMVGMWLMG 1146
Qy 1161 ISVNAVSLNLVSAVMSVEFVSHITRSPAISTKPTWLERAKEATTSMGSAPVAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVSKGSRVERAEALAHMGSVFSGITLTK 1206
Qy 1221 LPEGLVLGAKAQLIQIFFRNLMLTLTLGLHLGLVFLPVILSVYGPVNPA 1272
Db 1207 FGSIVVLAFAKSQIFQIFVFRMYLAMVLLGATHGLIFLPVLLSYIGPSVYKA 1258

RESULT 72

US-11-222-021-441
; Sequence 441, Application US/11222021
; GENERAL INFORMATION:
; APPLICANT: Bruno DOWON et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001548
; CURRENT APPLICATION NUMBER: US/11/222,021
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 7554
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 441
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
US-11-222-021-441

Query Match 34.8%; Score 2402.5; DB 42; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
Qy 7 RGWLLWALLRLRLAQSEPYTTIHQPGYCAFYDECG-----KNPELSGSLMTLSNVSCLS 59

Db 4 RGLALGLLLLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEYSG----- 46
Qy 60 NTPARKITGDHILILLOKICPRLYTPGNTQACCSAKQLVLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLOTLKDNLOLPQLQFLSRCPSCFYN 103
Qy 120 FWNLHCHMTCPNOSLFINVTR---VAQLGAGQLPAAVVAEAFYQHSFAEQSYDSCSRV 175
Db 104 LLNLECELTCPRQSQOFLNVTATEDYVDVNTQNTKNKELOYYYVQGSFANAMYNACRDV 163
Qy 176 RVPAAATLAVGTMCVYGSALCNAQRWLNFOGDTGNGLAPLDIT-----PHLLEPGQA 228
Db 164 EAPSSNDKALGLLCKGKADA-CNATNWIETMFKNDNGQAPFTTTPVFSDFPVH----- 215
Qy 229 VGSQTQPLNEGVARCNESQGDVATCQDCQCAASCPAATRPQ-----ALDSTFYLG 279
Db 216 ---GMEPNNATKGCDESDEVTAFCSCQDCSIVCGPKPQPPPPAPMTILGDAMYVIM 272
Qy 280 QMPGSLVLIILCSVPAAVVTTILL-----VGRVAPARDKSMVDPKKGTSLS 326
Db 273 WITYMAFLVFPFAFVAVWCYRKRYFVSEYPTIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLGQFFQCGMTWASWPLTILVLSVIPVVALAAGLVFTELTTPDVELWSAP 386
Db 325 DPVSAAPBGLRLRFPTRGSCFCVRNPGCVIPFSLVFTACSSGLVFRVTTNPVDLWSAP 384
Qy 387 NSQARSEKAFHDQHPFRFTNOVLTAPNRSSRYDSSLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDHQFPFTEQLIRAPLTDKHTYQYPVPGADVPFGPPLDIQLHLQVL 444
Qy 446 ELQERLRHLQVMSPEAQRNISLODI CYAPLNPDNLSYDCCINSLOVFNQNRITLLTA 505
Db 445 DLQIAIEN--ITASVDNETVLODILCLAPLSYNT---NCTILSVLNFQNSHSLDHHK 499
Qy 506 NQTLMGQTSQVMDKHDFLYCANAPLTKDGTALALS CMADYGA PVPFPLAIGYKGKDY 565
Db 500 GDDPF---VYADYHTFLYCVRAPASLNDTSLHDPCLGTGCGPVFPMLVLGGYDQNYN 556
Qy 566 EEAALIMTFSLNYPAGDPRLAQAKLWEEAFLEEMRAFORMAGMFOVTFTEARSLDEI 625
Db 557 NATALVITPPVNNYNDTEKQRAQAWKEKFNFVKNYKN---PNLTISFTAERSTIEDEL 613
Qy 626 NRTTAEDLPFATSVIVIFLYISLALGSYSSSRVMDSKATLGLGGVAVVLGAYMAAMG 685
Db 614 NRESDSVFTVVISVAINFELYSIALGHKSCERLLVDSKVSGLTAGILVLSSVACSIG 673
Qy 686 FFSYLGIIRSLVILQVPPFLVSVGADNIFIFVLEYQRLPRPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVLPFLVAVGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
Qy 746 SMLCSLSEAICFFLGALTPMPAVTFTALTSGLAVIDLFLQMSAFVALLSDSKROEAS 805
Db 734 SMLFSFSETVAFILGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIKRQEN 793
Qy 806 RLDVCCVKPQELPPPGQ-GEGLLIGFFOKAVAPFLHWHITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVGAEDGTSVQASESCLFRFFKNSVSPLLDKDMPPIVIAIFVGLSFSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNPFSSAGNAICSSAG 924
Db 854 NKVDIGLDQSLMPDSDSYMDVYFKSISQYLHAGPPVYFVLEGRHDYTSKGGQNMVCGG 913
Qy 925 CNPFSTQKIQYATEFPEQSYLAIAPASSWDDFIDWLT-SSCCRLYISGPNKDECPST 983
Db 914 CNNDSLVQOIFNAALQDNTRYIGFAPSSWIDDYFDWVKPQSSCCRY---DNITDQFCNAS 970
Qy 984 VNSLNLCKNMSIT-MGSVRPSVEOFHKYLPFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDDPACVR-CRPLETPEGKORPOGGDFMRFLPFLSDNPNPKCKGGAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFPPYITNV 1100

Db 1030 GTRVGATYFMTYHTVLTQTSADFDALKKARLIASNT-ETMGINGS--AYRVPYSVYV 1086
Qy 1101 FYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFMAWMD 1160
Db 1087 FYEQYLTIDDTITFNLGVSIGAFVLTWMLLGCWLMVSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLNLVAGVMSVEFVSHITRSPASTKPTWLBRAKATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKSRVERABEALAHMSSVPSGTLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFPFRMLNLLITLGLLHGLVFLPVLVSYVGDVNP 1272
Db 1207 FCGIVVLAPAKSQIQIFYFRMYLAWVLGATHGLIFLPLVLSYIGPSVYKA 1258
RESULT 73
US-60-423-552-148
; Sequence 148, Application US/60423552
; GENERAL INFORMATION:
; APPLICANT: American Home Product Inc.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING MAJOR DEPRES
; FILE REFERENCE: AM101250L
; CURRENT APPLICATION NUMBER: US/60/423,552
; CURRENT FILING DATE: 2002-11-05
; NUMBER OF SEQ ID NOS: 256
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 148
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-423-552-148
Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
Qy 7 RGWLLWALLRLAQSPPYTHIOPGYCAFVDEG-----KNPELSSGLMTLSNVCSLS 59
Db 4 RGIALGLLLLLCPAQVFSQ-----SCVMYGECCYAGDKVNCYSG----- 46
Qy 60 NTPARKITGDHLLILLOKICPRLTYGNTQACCSAKOLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGVLDVQLCPGPFPG-NVSLCCDVRQLTKNLQNLQPLQFLSRCPSCFTN 103
Qy 120 FVNLHCNTCSNOSLFINVTR----VAOLGAGQLPAVVAEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPQSOFLNWTATEDYVDPVNTQKTNVRELQYYVQGSFANMAMNACRDV 163
Qy 176 RVPAAATLAVGTMCYGYGSALCNAORWLNFGQDTGNGLAPLDIT-----FHLLEPGQA 228
Db 164 BAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTITPVSPDPVH----- 215
Qy 229 VSGIOPLENGVARNESOGDDVATCSQDCOAACPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDSVDEVTAPGSCQDCSIVCGPKQPPPPAPWPTILGIDAMYIM 272
Qy 280 QMPGSLVLLILCSFAVVITLL-----VGFVAPARDKSKWDPKGTSL 326
Db 273 WITYMAFLVFFGAPFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLGQFQCGWGTWASWPLTILVLSIPWVLAAGLVFTLTTPDVELWSAP 386
Db 325 DPVSAFAFECCLRLFRMGSFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
Qy 387 NSQARSEKAFHOFQFPFRTRNQVILTAPNRSYRVDLSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDFQFGFPFRTEQLIIRAPLTDKHIYQPYPSGADVPFGPDLQILHVL 444
Qy 446 ELQERLRLQVMSPEAQRNLSLODICVAPLNPDNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLODICLAPLSYNT---NCTILSVNLVFNQNSHVDHKK 499

Qy 506 NQTLMGQTSQVDKDHFLYCANAPLTFKGTALALASCMADYAGAPVFPFLAIGYKGDYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLDPCLTGTFGGVFPFWLVGYYDQNYN 556
Qy 566 EBAELMTSLNNYPAGDPRLAQAKLEBAFLIEMRAFORRMAGMFMQVPTTASRSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKLQBAQAWKEFNFVKNYK---PNLTISTASRSLEDEL 613
Qy 626 NRTAEDLPFPATSYIVIFLYISLALGSSSRVMDVKATGLGAGVAVVLGAVMAAMG 685
Db 614 NRESDSDFVTWISYAIMFLYISLALGHSKRRLLVDSKVSIGIAGILVLSVACSLG 673
Qy 686 PFSYLGISSVLIVQVVPFLVSVGADNIFIFVLYEYORLPRRPGEPREVIHGALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFELVLAQVNDIFILVQAYQDRERLQGETLDOOLGRVLGEVAP 733
Qy 746 SMLLCSLSAICFFLICALTPMPAVRTFALTSGLAVIDLDFLOMSAFVALLSLDSKQREAS 805
Db 734 SMFLSSFTVAFPLGALSVMFAVHTFSLPAGLAVFIDFLQLTCTFVSLGLDIIKQEK 793
Qy 806 RLDVCCVQPOBLPPPGQ--GEGILLGFQKAYAPFLHMTITRGVLLLLFLALFVSLYS 864
Db 794 RLDIFCCVGAEDGTSVQASECLPFFKNSYSPLLKDMRPIVIAIFVGVLSFSAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLNRYFVPGAPYVFTTGLYFNSSSEAGMAICSSAG 924
Db 854 NKVDIGLDSMPDDSYMVDYFKSISQYLHAGPPVYFVLEEGHDYVTSKQGNVCGMG 913
Qy 925 CNPFSTOKIOYATEPPEQSILAI PASSWVDDFIDMLTP--SSCRLYISGPNKDKCPST 983
Db 914 CNNDLSVQOIFNAAQLDNTRYIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQPCNAS 970
Qy 984 VNSLNLCKNCSIT--MGSVRPSVEQHFYLPFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR--CRPLTEGKQRPQGGDFMRFLPFLSDNPNPKCGKGGHAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAVHKJNSQDYTEALRAARELAANTADLRKVGTDPAPFEVPPYITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADFDALKKARLIASNT-ETMGINGS--AYRVPYSVYV 1086
Qy 1101 FYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIWMLVDTVGFMAWMD 1160
Db 1087 FYEQYLTIDDTITFNLGVSIGAFVLTWMLLGCWLMVSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLNLVAGVMSVEFVSHITRSPASTKPTWLBRAKATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKSRVERABEALAHMSSVPSGTLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFPFRMLNLLITLGLLHGLVFLPVLVSYVGDVNP 1272
Db 1207 FCGIVVLAPAKSQIQIFYFRMYLAWVLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 74

US-60-427-579-148
; Sequence 148, Application US/60427579

; GENERAL INFORMATION:

; APPLICANT: American Home Product Inc.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING MAJOR DEPRE

; FILE REFERENCE: AM101250L

; CURRENT APPLICATION NUMBER: US/60/427,579

; CURRENT FILING DATE: 2002-11-20

; NUMBER OF SEQ ID NOS: 256

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 148

; LENGTH: 1278

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-427-579-148

Query Match 34.8%; Score 2402.5; DB 48; Length 1278;

Best Local Similarity 39.8%; Pred. No. 2.3e-201;


```
QY 446 ELQERLRHLQVWSPQAQRNLSLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVNLVFNQSHSVLDHKK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKQGTALALSCMADYDGPVPPFLAIGYKGYK 565
Db 500 GDDFF---VYADYHTEFLYCVAPASLNTSLHDFCLGTFGGVPFWLVGLGYDDQYN 556
QY 566 EBAALIMTSLNYPAGDPRLLAQAKLWEAFLEEMRAFORRMAGMPQVTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAWKEBFNFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYTSLALGSSVSSRSRVMDSKATLGLGVAVVLGAVMAWG 685
Db 614 NRESDSDFVTVVISYAIMFLYISLALGHKSCRRLLVDSKVSILGILVILSVVACSILG 673
QY 686 FFSYLGRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
QY 746 SMLLCSLSEAICTFFLGALTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKREAS 805
Db 734 SMFLSSFSFETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIREQEN 793
QY 806 RLDVCCCVKQBLPPPGQ--GEGLLIGFPQKAYAPFLLHWITRGVULLFLFALFGVLSYLM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNYSPLLLKDMWRPIVIAIFVGLSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFVGPVYFVTTLYGNFSSSEAGNAICSSAG 924
Db 854 NKVDIGLQDSLMPDSDSYMDYFYSISQYLHAGPPVYFVLEBHDYTSKQGNMVCGMG 913
QY 925 CNNFSTQKIQYATEPPEQSQYLAIAPASSWVDDFDWLTP--SSCCRLYISGPNKDFCPS 983
Db 914 CNNDLSVQIQIFNAQOLDNYTRIGFAPSSWIDDYFDMVKQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT--MGSVRPSVEOPHYKLPWFLNDRPNIKCPKGLAAYTSVN--LFS 1040
Db 971 VVDPACVR--CRPLTPEGKQRPQGGFMRLPMLFSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQOVLASRFMAWKPLKNSQDYTEALRAARELANITADLRKVGTDPRAFEVFTITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADIDALKKARLASNT--ETWINGS--AYRFPYSVYV 1086
QY 1101 FVEQYITLPEGLFMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FVEQYITLIDDTIFNLGVSGLAIFLVTVLLGCELWSAIVMCAITANVLVNMFGVWLWG 1146
QY 1161 ISYNAVSLNLVSAGMSVEFVSHITRSPAISTKPTWLERAKETISMGSAYFAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSGSGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLTK 1206
QY 1221 LPEGLVGLAKAQLQIOPFRNLTLITLGLLHGLVFLVILSVYGPVWPA 1272
Db 1207 FGGIVVLAFAKSOIQIFRYRMYLAMVLLGATHGLIFLVLVSLSYGSPVKA 1258

RESULT 76
US-60-453-135-14124
; Sequence 14124, Application US/60453135
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001456
; CURRENT FILING DATE: 2003-03-10
; CURRENT FILING DATE: 2003-03-10
; NUMBER OF SEQ ID NOS: 82762
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14124
; LENGTH: 1278
; TYPE: PR1
```

```
; ORGANISM: Homo sapiens
US-60-453-135-14124
```

```
Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGLWALLRLRLAQSEPTTHIQPGYCAFYDEGC-----KNPELSSGLMTLSNVSCLS 59
Db 4 RGLALGLILLLLCPAQFSQ-----SCVMYEGCIAYGDKRYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFG--NVSLCCDVRQLTKDNLQLPQFLSRCPSCFYN 103
QY 120 FVNLHCNTCSNQSFLINVT---VAOLGAGQLPAVVAIFYQHSFAEQYSQSCSRV 175
Db 104 LNLFLCFLTCSPQSQFLNVTATEDYDVPVNTQTKNVKELQYYVQSQSPANMYNACRDV 163
QY 176 RYPAATLAVGTMCYVYGSALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLBPGQA 228
Db 164 EAPSNDRKALGLLCKGKADA--CNATNWIEMFNKONGQAPFITTPVSPDPVH-----215
QY 229 VCSGIQPLNEGVARNCSQGDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNNAATKGCDESVDEVTAPCSCQDCSIVCGPKQPPPPAPAPWTILGLDAMVIM 272
QY 280 OMPGSLVILLILCSVPVAVTILL-----VGPRVAPARDKSKMVDPKKGTSL 326
Db 273 WITYMAFLVFPFAGFAVWCYRKRKFVSEYTPIDSNIAFSV--NASDKG-----ASCC 324
QY 327 DKLSFTHTLLQFFQGGWGTWASWPLTLLVLSVIPVVALAAGLVTELTTPDVELWSAP 386
Db 325 DPVSAAFEGCLRLTRWGSFCVRNPGCVIFPSLVPITACSSGLVFRVTTPVVDLWSAP 384
QY 387 NSQARSEKAFHDQHPFPRTNQVILTAPNRSYRYSLLGLPK--NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDHFQFPFTEQLIIRAPLTKHIYQYPYSGADVFPFGPDLQILQVL 444
QY 446 ELQERLRHLQVWSPQAQRNLSLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICIAPLSYNT---NCTILSVNLVFNQSHSVLDHKK 499
QY 506 NOTLMGQTSQVMDKDFLYCANAPLTFKQGTALALSCMADYDGPVPPFLAIGYKGYK 565
Db 500 GDDFF---VYADYHTEFLYCVAPASLNTSLHDFCLGTFGGVPFWLVGLGYDDQYN 556
QY 566 EBAALIMTSLNYPAGDPRLLAQAKLWEAFLEEMRAFORRMAGMPQVTTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAWKEBFNFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYTSLALGSSVSSRSRVMDSKATLGLGVAVVLGAVMAWG 685
Db 614 NRESDSDFVTVVISYAIMFLYISLALGHKSCRRLLVDSKVSILGILVILSVVACSILG 673
QY 686 FFSYLGRSLVILQVVPFLVSVGADNIFIFVLEYQRLPRRGPBPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVAVGVDNIFILVQAYQDERLQGETLDQOLGRVLGEVAP 733
QY 746 SMLLCSLSEAICTFFLGALTPMPAVRTFALTSGLAVIDLDFLLQMSAFVALLSLDSKREAS 805
Db 734 SMFLSSFSFETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDIREQEN 793
QY 806 RLDVCCCVKQBLPPPGQ--GEGLLIGFPQKAYAPFLLHWITRGVULLFLFALFGVLSYLM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNYSPLLLKDMWRPIVIAIFVGLSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLNRYFVGPVYFVTTLYGNFSSSEAGNAICSSAG 924
Db 854 NKVDIGLQDSLMPDSDSYMDYFYSISQYLHAGPPVYFVLEBHDYTSKQGNMVCGMG 913
QY 925 CNNFSTQKIQYATEPPEQSQYLAIAPASSWVDDFDWLTP--SSCCRLYISGPNKDFCPS 983
Db 914 CNNDLSVQIQIFNAQOLDNYTRIGFAPSSWIDDYFDMVKQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT--MGSVRPSVEOPHYKLPWFLNDRPNIKCPKGLAAYTSVN--LFS 1040
Db 971 VVDPACVR--CRPLTPEGKQRPQGGFMRLPMLFSDNPNPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQOVLASRFMAWKPLKNSQDYTEALRAARELANITADLRKVGTDPRAFEVFTITNV 1100
Db 1030 GTRVGATYFMTYHTVLTQTSADIDALKKARLASNT--ETWINGS--AYRFPYSVYV 1086
QY 1101 FVEQYITLPEGLFMLSCLVPTFAVSCILLGLDLSGLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FVEQYITLIDDTIFNLGVSGLAIFLVTVLLGCELWSAIVMCAITANVLVNMFGVWLWG 1146
QY 1161 ISYNAVSLNLVSAGMSVEFVSHITRSPAISTKPTWLERAKETISMGSAYFAGVAMTN 1220
Db 1147 ISLNAVSLNLVMSGSGISVEFCSHITRAFTVSMKGSVRERAEALAHMGSSVFSGITLTK 1206
QY 1221 LPEGLVGLAKAQLQIOPFRNLTLITLGLLHGLVFLVILSVYGPVWPA 1272
Db 1207 FGGIVVLAFAKSOIQIFRYRMYLAMVLLGATHGLIFLVLVSLSYGSPVKA 1258
```

```
Db 914 CNNDLSVQOIFNAQAQNDNYTRIGFAPSPSWIDDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNCNLCNMSIT-MGSVRPSVEQFKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMFLPMFLSDNPNPKCGKGAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATYFMYHTVLTQTSADFDALKKARLIASVNT-ETWINGS--AYRVPYSVFYV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLSTVWILVDTVGMALWD 1160
Db 1087 FVEQYLTITDDTIFNLGSLGAIFLVTWVLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIFFRNLNLIITLGLLHGLVFLPVILSYVGPDPVNP 1272
Db 1207 FGGIVVLAFAKSQIQIFFRMYLAMVLGATHGLIFLPLVLSYIGPSVYKA 1258

RESULT 77
US-60-466-412-14124
; Sequence 14124, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14124
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-466-412-14124

Query Match 34.8%; Score 2402.5; DB 48; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLRLAQSEPYTHTHQPGYCAFYDECG-----KNPELSGLSMTLSNVCLSL 59
Db 4 RGLALGLLALLLCPAQVFSQ-----SCVMYEGCGIAYGDKRYNCEVSG-----46

Qy 60 NTPARKITGDHILLQKICPRLYTGNTQACCSAKQLVLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPLQLFSLRCPSCFYN 103

Qy 120 FYNLCHNTCSNOSLFINVTR-----VAQLGAGQLPAVVAEAFYQHSFAGSVDSCSRV 175
Db 104 LNLFLCELTCSPQSOFLNVTATEDYDVPVNTQTKNVKELQYYVQGSFANAMYNACRDV 163

Qy 176 RYPAATAATVAGTCMGVYGYSALCNAORWLNFGQDGTNGLAPLDIT-----PHLLEPGQA 228
Db 164 EAPSSNDKALGLLGGKQADA-CNATNWIEMFNKONGQAPFTITPVFSDFFPVH-----215

Qy 229 VGSGLQPLNEGVARCNESGGDDVATCSQDCAAASCPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNNAKGCDESDEVTAPCSQDCSIVCGPKPQPPPPAPWTLILGLDAMYIM 272

Qy 280 QMPGSLVLIILCSVFAVVTILL-----VGFVPAPARDKSKMVDPKKGTSL 326
Db 273 WITYWAFLLVFGAFPAVWCYKRYFVSEYTPIDSNIAFSV-NASDKG-----ASCC 324

Qy 327 DKLSFSTHTLLQGFQCGWGTWVASWPLTILVLSVPVVAALAGLVFTLTDPVLSWAP 386
Db 325 DPVSAFAEGLCLRLFRMGSFCVRNPGCVIFFLSVFITACSSGLVFPVRVTINPVDLWSAP 384
```

```
Qy 387 NSQARSEKAPHDQHFQFFERTNOVILTAPNRSSYRYSLLGPK-NFSGILLDLDLLL 445
Db 385 SSQARLEKEYDQHFQFFERTNOVILTAPNRSSYRYSLLGPK-NFSGILLDLDLLL 444
Qy 446 ELQBLRHLQWSPQAQRNLSQDICVAPLNPDNTSLYDCCINSLLQYFQNNRTLLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLQDICLAPLSPYNT--NCTILSVNLNFQNSHSLDHKK 499
Qy 506 NQTLMGQTSQVDKDHLYCANAPLTFKDGITALALSCHADYCAPVPPFLAIGGYKGDYS 565
Db 500 GDDEFF--VYADYHTHELYCYRAPASLNDTSLLDHDPCLGTGPGVPFVPLVGGYDQNYN 556
Qy 566 EAEALIMTFSENNYPAGDPRLAQAALMEEAFLEEMRAFORRMAGMFOVTFEATERSLEDEI 625
Db 557 NATVALVTFPNNYINDEKQRAQAKEKEFINFKVKNY---PNLTISFTAESRIDEEL 613
Qy 626 NRTTAEDLIPATSYIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAYMAAAG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHIKSCRLLVDSKVSGLTAGILVLSSVACSLG 673
Qy 686 PFSVLGTRSSLVILQVVPFLVLSVAGADNIFIVLEYQRLPRRPGEPREHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIPFLVAVGVNDIFLVOAYQDERLQGETLDQGLGRVLEGPVAP 733
Qy 746 SMLLCSLSEATCFELGALTMPAVRTFALTSGLAVIDFLQMSAFVALLSLDSKQBAS 805
Db 734 SMFLSSFSETVAFFLGALUSVMPAVHTSLFAGLAVFIDFLQITCFVSELGLDIYKQKXN 793
Qy 806 RLDVCCCKVQELPPPGQ-GEGLLLGFFQKAYAPPLLHWITRGVVLVLLPLALFGVLSYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASSECLFRFPKNSYSPLLKDMWRPIVIAIFVGLVSPSIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFNRYFVGAPVYFVTTLGYNFSSEAGNAALCSSAG 924
Db 854 NKVDIGLDQSLSMFDDSYMDVDFKSIQYHLHAGPPVYFVLEEGHDYTSKQGNMVCYGGMG 913
Qy 925 CNNEFSTQKIQYATEFPEQSYLAIPASSWVDDFIDWLTLP-SSCCRLYISGPNKDFCPSST 983
Db 914 CNDSLVQOIFNAQAQNDNYTRIGFAPSPSWIDDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNCNLCNMSIT-MGSVRPSVEQFKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMFLPMFLSDNPNPKCGKGAAYSSAVNILLGH 1029
Qy 1041 DQOVLASRFMAHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATYFMYHTVLTQTSADFDALKKARLIASVNT-ETWINGS--AYRVPYSVFYV 1086
Qy 1101 FVEQYLTILPEGLFMLSCLVPTFAVSCILLGLDLRSGLLNLSTVWILVDTVGMALWD 1160
Db 1087 FVEQYLTITDDTIFNLGSLGAIFLVTWVLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSVERAEEALAHMGSSVFSGITLTK 1206
Qy 1221 LPGAIVLGLAKAQLIQIFFRNLNLIITLGLLHGLVFLPVILSYVGPDPVNP 1272
Db 1207 FGGIVVLAFAKSQIQIFFRMYLAMVLGATHGLIFLPLVLSYIGPSVYKA 1258
```

RESULT 78

US-60-524-882-700

; Sequence 700, Application US/60524882

; GENERAL INFORMATION:

; APPLICANT: CARGILL, Michele

; APPLICANT: IAKOUBOVA, Olga

; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

; FILE REFERENCE: CL001494

; CURRENT APPLICATION NUMBER: US/60/524,882

```
; CURRENT FILING DATE: 2003-11-26
; NUMBER OF SEQ ID NOS: 46672
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 700
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-524-882-700

Query Match      34.8%; Score 2402.5; DB 49; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGLWLLALLRLRLAQSEPTTHIQPGYCAPYDECG-----KNPELSGLMTLSNVCSLS 59
D 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYDGRKYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
D 47 --PPKPLPKDGVLDVQELCPGFFFG--NVSLCCDVRQLQTLKDNQLQPLQFLSRCPCSFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAPQSDYSCSRV 175
D 104 LNLFLCELTCSPRQSQFLNVTATEDYDVPVNTQTKNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCVGYGALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGQA 228
D 164 EAPSSNDKALGLLCKGDADA-CNATNWIETMFKNDNGOAPFTITPVFSDFPVH-----215
QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCQCAASCPAIPARQ-----ALDSTFYLG 279
D 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
D 273 WITYNAPLVPFGAPFAVWCYKRYFVSYYTIDNSIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLQOFGQGTWVWASWPLTILVLVSVVVALAAGLVFTLTDPVBLWSAP 386
D 325 DPVSAAFEGCLRLFRWGSFCRPNPGCVFVPSLFTACSSGLVFRVTTNPVLDWSAP 384
QY 387 NSQARSEKAFPHQHGPPRTHQVILTAHRSYRYSLLGLPK-NFSGILDLLELL 445
D 385 SSQARLEKEYFQHFQFPFTEQLIIRAPLTDKHIYQYPFGADVPFGPPLDIQLHVL 444
QY 446 ELQERLRHLQVMSPEAQRNLSQDICYAPLNDPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
D 445 DLQIAIEN--ITASVDNETVTIQQICLAPLSYNT---NCTLSLVNYPQNSHSLDHKK 499
QY 506 NOTLMGQTSQVDWKDHFLLYCANAPLTFKDGITALALS CMADYGAPVFPFLAIGYKGDYS 565
D 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLHDPCLGTGFGPVPVPLVLGSDYDQNVN 556
QY 566 EBAALIMTSLNYPAGDPRLAQAKLEBAFLBEMRAFQRMAGMFPQVTFPAAERSLEDEI 625
D 557 NATALVITPVNNYNDTEKLQRAQAKEKEFTNFVKYKN---PNLITISFTAERSIDEI 613
QY 626 NETTAEDLPFATSYIVIFLYLSLALGSYSSRWSRVWDSKATLGLGGVAVVLGAWAAG 685
D 614 NRESQSDVPTVVISVAIMFLYISLALGHITKSCRRLLVDSKSLGAGILVLVSSVACSIG 673
QY 686 FFSYLGISSVLIVQVPLVLSVADNIFIFVLEYQRLPRPGEPREVIHGRALGRVAP 745
D 674 VFSYIGLPLTLVIEVPLVAVGVNDIFILVQAYQDERLOGETLQDQGLGRVLEVAP 733
QY 746 SMLLCSLSRAICFFIGALTPMPAVRTFALTGLAVILDFLQMSAFVALLSDSKRQAS 805
D 734 SMFLSSFSFETVAFFLGALSVMBAVHTFSLFAGLAFIDFLQITCFVSLGLDIDKQKKN 793
QY 806 RLDDVCCVCKPQLPPPGQ--GEGLLIGFFQKAYAPFLLHWTIGVVLVLLFLALFGVLSYM 864
D 794 RLDFCCVRGABDGTSTVQASESLFRFPFKNSYSPLLKDKMMRPVIAIPVGVLSFIAVL 853
```

```
QY 865 CHISVGLDQELALPKDSYLLDYPLFLNRYFEVGA PVYFTVTLGYNFPSSBAGNAICSSAG 924
D 854 NKVDIGLDQSLSPDDSYMVDYFKSISQYLHAGPPVYFVLEBCHDYTSKGGQNMVCGMG 913
QY 925 CNPFSTOKIQVATEPPEQSYLAIPASSWVDDFIDWLTG-SSCCRLYISGPNKDKPCST 983
D 914 CNNDLSVQOIFNAQOLDNYTRIGFAPSPSSWIDDYFDWVFPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCLNCSMIT-MGSVRPSPVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
D 971 VDPACVR-CRPLTEBGRQPOGGDPMRPLPMLFSDNPNPKCGKGGHAAYSANVILLGH 1029
QY 1041 DQCVLASRPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAPFVPPYITNV 1100
D 1030 GTRVGATFMTYHTVLTQTSADPIDALKKARLTASNT-ETMGINGS--AYRVFPYSVYV 1086
QY 1101 FYEQVLTITLPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIVMILAVDTVGMALWD 1160
D 1087 FYEQVLTITIDDTIFNLGVSGLAIFLVTWVLLGCELSAVIMCATIAMLVNMFVNMWLG 1146
QY 1161 ISYNAVSLINLVSAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVPAVAMTN 1220
D 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRVERAEALAHMGSSVFSGITLTK 1206
QY 1221 LPGAIVLGLAKAQLIOIFFRNLNLTLLGLLHGLVFLPVILSYVGPVNP 1272
D 1207 FGGIVVLAFARQIQIPIFFFRMYLANVLLGATHGLIFLPLVLSYIGPSVKA 1258
```

RESULT 79

```
US-60-568-219-476
; Sequence 476, Application US/60568219
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: RESPONSIVENESS TO STATIN TREATMENT, METHODS OF DETECTION AND
; FILE REFERENCE: CL001516
; CURRENT APPLICATION NUMBER: US/60/568,219
; CURRENT FILING DATE: 2004-05-06
; NUMBER OF SEQ ID NOS: 28217
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 476
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-568-219-476
```

```
Query Match      34.8%; Score 2402.5; DB 49; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGLWLLALLRLRLAQSEPTTHIQPGYCAPYDECG-----KNPELSGLMTLSNVCSLS 59
D 4 RGLALGLLLLLCPAQVFSQ-----SCVMYGECAIYDGRKYNCEYSG-----46
QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
D 47 --PPKPLPKDGVLDVQELCPGFFFG--NVSLCCDVRQLQTLKDNQLQPLQFLSRCPCSFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFAPQSDYSCSRV 175
D 104 LNLFLCELTCSPRQSQFLNVTATEDYDVPVNTQTKNVKELQYVYGQSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMCVGYGALCNAQRWLNFGQDGTGNGLAPLDIT-----PHLLEPGQA 228
D 164 EAPSSNDKALGLLCKGDADA-CNATNWIETMFKNDNGOAPFTITPVFSDFPVH-----215
QY 229 VGSGLQPLNEGVARCNESQGDVATCSQDCQCAASCPAIPARQ-----ALDSTFYLG 279
D 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKPQPPPPAPMTILGDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVVTILL-----VGRVAPARDKSMVDPKKGTSLS 326
```

```
Db      273 WITYMAFLVFGCAFAVWCYKRYFVSYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy      327 DKLSFSTHLLGQFFGCGGTWASWPLTILVSVIPVVALAAGLVFTLTTPDVELWSAP 386
Db      325 DPVSAAFEGCLRLFRWGSFCVRNPGCVIFFSLVFITACSSGLFVRVTINPVDLWSAP 384
Qy      387 NSQARSEKAFDQHGFPFRTNQVILTAPNRSYRYSLSLLGPK-NFSGILDLLLELL 445
Db      385 SSQARLEKEYFDQHGFPFRTQOLIIRAPLTKHIIQYPYPSGADVPFPGPLDIQILHQVL 444
Qy      446 ELQERLRHLQVMSPEAQRNISLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db      445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT--NCTILSVLNYFQNSHVDLHKX 499
Qy      506 NOTLMGQTSQVWDKHFHLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGYKGDYS 565
Db      500 GDDPF---VYADYHTHFLYCVRAPASLNDTSLLDHDPCLGTGCGPVFPMVLGGYDQNYN 556
Qy      566 EAEALIMTFSLANNYPAGDPRLAQAKLWEAEFLBEMRAFORRMAGMFPQVTFTAERSLEDEI 625
Db      557 NATALVITFPVNNYNDTEKLOQAQWKEKFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy      626 NRTAEDLPFATSYVIVFLYISLALGSYSWSRVMVDSKATILGCGVAVVLGAVMAAMG 685
Db      614 NRESDSVFTVTVISVAINFLYISLALGHIKSCRLLVDKSVLSGIAGIILVLSVACSIG 673
Qy      686 PFSYLGIIRSSILVQVVPFLVSYGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAP 745
Db      674 VFSYIGLPLTLVIEVIPFLVAVGVNDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
Qy      746 SMLLCSLSBAICFFLICALTPMAVRTFALTSLGLAVILDLLQMSAFVALLSLSKQRAS 805
Db      734 SMFLSFSSTVAFGLGALSVMAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIIRQKKN 793
Qy      806 RLDVCCVKBQELPPGQ--GEGLLGFFQKAYAPFLHWHITRGVLLFLFALFGVSLYSX 864
Db      794 RLDIFCCVGAEDGTSVQASESCLPRFPKNSYSPILLKXDMRPVIAIFVGLVSFI AVL 853
Qy      865 CHISVGLDELALPKDSYLLDFLFLNRYFVAGAPVYFTTLYGNFSSBAGNATCSSAG 924
Db      854 NKVDYGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYTSKGGQNMVCGG 913
Qy      925 CNNFSFTQIQVATEPPEOSYLAIAPASSWVDDFDWLTP--SSCCRLYISGPNKDKPCST 983
Db      914 CNNDLSVQOIFNAQLNDYTRIGFAPSSWIDDYFDWVKFQSSCCRV---DNITDQFCNAS 970
Qy      984 VNSLNLCKNCSIT-MGSVRPSVEQPHKYLPHWFLNDRPNIKCPKGGLAAYSTSVN--LTS 1040
Db      971 VVDPACVR--CRPLTPEGKQPOGGDFMFLPMLFSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy      1041 DGOVLASRFMAVHKPLKNSODYTEALRAARELANITADLRKVPGTDPAPFVFPYTIINV 1100
Db      1030 GTRVGATYFMTVHTVLTQTSADFDLAKKARLASNT-ETWINGS--AIRVFPYSVFV 1086
Qy      1101 FYEQYLITLPEGLFMLSCLVPTFAVSCILLGLDLRSGLNLLSVMLIVDVTGFWALWD 1160
Db      1087 FYEQYLITIDDTIFNLGVSGLGALFLVTWVLLGCELWSAVIMCATAMVLVNMFGVNLWG 1146
Qy      1161 ISYNAVSLNLVAVGMSVEFVSHITRSPAITKPTWLERAKEATISMGSAVFAVAMTN 1220
Db      1147 ISLNAVSLNLVMSCGISVEFCSHITRAFTVMKGSVRVERAEALAHMGSSVFSGTILTK 1206
Qy      1221 LQGIVLGAKAQOLQIIFPRNLTLTLGLLGHGLVFLPVILSYVGPQWNP 1272
Db      1207 FGIIVVLAFAKSGIPIQIFRYRMLAMVLGLGATHGLIFLPLVLLSYIGPSVNKA 1258
```

RESULT 80

US-60-592-195-129

; Sequence 129, Application US/60592195

; GENERAL INFORMATION:

; APPLICANT: Applera Corporation

```
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01536PROV
; CURRENT APPLICATION NUMBER: US/60/592.195
; CURRENT FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 1278
; TYPE: PRF
; ORGANISM: Homo sapiens
; US-60-592-195-129
```

```
Query Match      34.8%; Score 2402.5; DB 49; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
Qy      7 RGWLLIALLRLAQSSEPYTTIHQPCYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
Db      4 RGLALGILLLLLLCPAQVFSQ-----SCVWYCECGIAYGDKRYNCEYSG----- 46
Qy      60 NTPARKITGDHLILLOKICPRLYTGPNTOACCSAKQLVSLSEASLITKALLTRCPACSDN 119
Db      47 --PKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDNLQLPLQFLSRCPSCFYN 103
Qy      120 FVNLHCHNTCSNOSLFINVTR---VAQLGAGQLPAVAYEAFYQHSFASQSYDCSRV 175
Db      104 LNLNFCELTCSPROSQFLNVRATEDYDPVTNQTKTNVKELOYYVGVGQFANAMYNACRDV 163
Qy      176 RVPAAATLAVGTTCGVYGSALCNAQRWLNFGDGTGNGLAPLDIT-----PHLLEPGA 228
Db      164 EAPSNDRKALGLCGKDADA--CNATNWLEYMFKNDNGOAPFTITPVFSDFPVH----- 215
Qy      229 VGSGLQPLNEGVARCNESQGDVATCSQDCAASCAPAIARPO-----ALDSTFVLG 279
Db      216 ---GMPEMNNATKGDSEVDVETAPCSQDCSIVCGPKPQPPPPPPAPWTILGLDAMYIM 272
Qy      280 QMPGSLVLIILCSVFAVVVITLL-----VGRVAPARDKSMVDPKKGTSLS 326
Db      273 WITYMAFLVFGCAFAVWCYKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy      327 DKLSFSTHLLGQFFGCGGTWASWPLTILVSVIPVVALAAGLVFTLTTPDVELWSAP 386
Db      325 DPVSAAFEGCLRLFRWGSFCVRNPGCVIFFSLVFITACSSGLFVRVTINPVDLWSAP 384
Qy      387 NSQARSEKAFDQHGFPFRTNQVILTAPNRSYRYSLSLLGPK-NFSGILDLLLELL 445
Db      385 SSQARLEKEYFDQHGFPFRTQOLIIRAPLTKHIIQYPYPSGADVPFPGPLDIQILHQVL 444
Qy      446 ELQERLRHLQVMSPEAQRNISLQDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db      445 DLQIAIEN--ITASVDNETVTLQDCLAPLSPYNT--NCTILSVLNYFQNSHVDLHKX 499
Qy      506 NOTLMGQTSQVWDKHFHLYCANAPLTFKGTALALSCMADYGAPVFPFLAIGYKGDYS 565
Db      500 GDDPF---VYADYHTHFLYCVRAPASLNDTSLLDHDPCLGTGCGPVFPMVLGGYDQNYN 556
Qy      566 EAEALIMTFSLANNYPAGDPRLAQAKLWEAEFLBEMRAFORRMAGMFPQVTFTAERSLEDEI 625
Db      557 NATALVITFPVNNYNDTEKLOQAQWKEKFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy      626 NRTAEDLPFATSYVIVFLYISLALGSYSWSRVMVDSKATILGCGVAVVLGAVMAAMG 685
Db      614 NRESDSVFTVTVISVAINFLYISLALGHIKSCRLLVDKSVLSGIAGIILVLSVACSIG 673
Qy      686 PFSYLGIIRSSILVQVVPFLVSYGADNIFIVLEYQRLPRRPGEPREVIHGRALGRVAP 745
Db      674 VFSYIGLPLTLVIEVIPFLVAVGVNDNIFILVQAYQDERLQGETLDOQLGRVLGEVAP 733
Qy      746 SMLLCSLSBAICFFLICALTPMAVRTFALTSLGLAVILDLLQMSAFVALLSLSKQRAS 805
Db      734 SMFLSFSSTVAFGLGALSVMAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIIRQKKN 793
```



```
; Sequence 197, Application US/60651235
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001581
; CURRENT APPLICATION NUMBER: US/60/651,235
; CURRENT FILING DATE: 2005-02-10
; NUMBER OF SEQ ID NOS: 3512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 197
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-651-235-197

Query Match      34.8%; Score 2402.5; DB 50; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLALLRLLAQSPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLALGULLLLCPAQVFSQ-----SCVWYGECCYAGDKRYNCEYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKONLQPLQFLSRCSCFYN 103

QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFASQSDYSCSRV 175
Db 104 LNLFLCELTCSQRSQSLNVTATEDYDVPVNTQKTNVKEQLYVVGQSFSANAMYNACRDV 163

QY 176 RVPAAATLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDT-----PHLEPQGA 228
Db 164 EAPSSNDKALGLCCGKADA--CNATNWIETMFKNGQAPFTITPVFSDFPVH-----215

QY 229 VSGIOPLNEGVARCNESQGDVATCSQDCAASCPAIRPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESDEVTAPCSCDCSIVCGPKQPPPPAPWTLILGDAMYIM 272

QY 280 QMPGSLVLIILCSFVAVTILL-----VGRFVAPARDKSRMDPKGTSLS 326
Db 273 WITYMAFLVFFGAFFVWCYRKRYFVSYPIDSNIAFSV--NASDKG-----ASCC 324

QY 327 DKLSFSTHTLLGFOFGWGTWASWELTILVSVIPVVALAAGLVFTELTDPVLSWAP 386
Db 325 DPVSAAFEGCLRLFTWGSFCVRNCGVIFSLVFITACSGLVFVRVITTFVLDWSAP 384

QY 387 NSQARSEKAFHDQHGPFRTNQVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDQHGPFRTNQVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 444

QY 446 ELQERLRLHQVMSPEAQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLTA 505
Db 445 DIQIALEN--ITASYDNETVTLQDICALPLSYNT---NCTLSLVNYPQNSHSDHKK 499

QY 506 NOTLMQOTSQVDWKDHFYCANAPLTFKDGTLALSCMADYGPVFPFLAIGSGYKDY 565
Db 500 GDDFF---VYADYHTEFLCYVRAPASLNDTSLHDPCLTGFGPFPVPLVGGYDDQYN 556

QY 566 EBAELIMTSLNYPAGDPRLLAQAKLWEAFLEEMRAFQRMFMQVFTTFAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEBFINFKVKN---PNLTISFTAERSIEDEL 613

QY 626 NRTABDLPIFATSYIVIFLYISLALGSYSSRSRVVDSKATLGLGGVAVVLGAVWAMG 685
Db 614 NRESUSDVTVISVIAIMFYISLALGHYKSCRLLVDSKISGLIAGILVLSVACSLG 673

QY 686 FFSYLGIRSSVLQVVPFLVLSVADNIFI VLEYQRLPRPFGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTVIEVPELVAVGVDNIFILVQAYQRDERLQGETLQQLGRVLGEVAP 733

QY 746 SMLLCSLSBAICFFIGALTMPMPAVTRFALTSGVLVLDPLLOKSAFVALLSDSKRQAS 805
Db 745 SMLLCSLSBAICFFIGALTMPMPAVTRFALTSGVLVLDPLLOKSAFVALLSDSKRQAS 805

Db 734 SMFLSSFSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQLQITCFVSLGLDIIKREQN 793
QY 806 RLDVCCCKVPQELPPGQ--GEGLLIGFFOKAYAPFLHWHITRGVLLLLFLALFGVSLYGM 864
Db 794 RLDIFCCVRGAEAGTGVQASESCLFRFFKNSYSPLLLKDMWRPIVIAIPGVLSFSIAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVAGPVYFVTTLGYNPFSSAGNAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMDVYFYSISQYLHAGPPVYFVLEGRHDYTSKQGNMVCGGMG 913
QY 925 CNNESTOKIOVATEFPEQSILAIIPASSWDDDFIDWLTP--SSCRLYISGPNKDKPCPST 983
Db 914 CNNDSLVQOIFVNAQOLDNYTRIGFAPSPSWIDDFWVRPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNCSIT--MGSVRPSVEQFHKYLPMFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR--CRPLTPEGKORPQGGDFMRFLPMFLSDNENPKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQCVLASRPMAYHKLKNSQDYTEALRAARELANITADLRKVPGTDPAFEPFYTITNV 1100
Db 1030 GTRVGATYFMTYHTVLOTSDAFIDALKKARLIASNT--ETMINGS--AYRFPYSVFPV 1086
QY 1101 FYEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRSGLLMLLSIVMLVDTVGFMAWD 1160
Db 1087 FYEQVLTIIDTIFNLGYSLGAIPLVTLVLLGCELWSAVIMCATIAMVLMVGMWMLMG 1146
QY 1161 ISYNAVSLINLSVAGMSVEFVSHITRSFAISTKTPTWLERAKEATISMSGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSGSIISVFCSHITRAFTVSMKGRVERAEEALAHMSSVFSGITLTK 1206
QY 1221 LPGTILVILGLAKAQLQIIFPRLNLLITLLGLLGLVFLPVLSYVGPDPNPA 1272
Db 1207 FGIIVVLAFAKSQLPQIFIFRMYLAMVLLGATHGLIFLPLVLLSYIGPSVNSKA 1258

RESULT 83
US-60-651-235-198
; Sequence 198, Application US/60651235
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001581
; CURRENT APPLICATION NUMBER: US/60/651,235
; CURRENT FILING DATE: 2005-02-10
; NUMBER OF SEQ ID NOS: 3512
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-651-235-198

Query Match      34.8%; Score 2402.5; DB 50; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLALLRLLAQSPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLALGULLLLCPAQVFSQ-----SCVWYGECCYAGDKRYNCEYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKONLQPLQFLSRCSCFYN 103

QY 120 FVNHLCHNTCSNQSLFINVTR-----VAQLGAGQLPAVVAYEAFYQHSFASQSDYSCSRV 175
Db 104 LNLFLCELTCSQRSQSLNVTATEDYDVPVNTQKTNVKEQLYVVGQSFSANAMYNACRDV 163

QY 176 RVPAAATLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDT-----PHLEPQGA 228
Db 164 EAPSSNDKALGLCCGKADA--CNATNWIETMFKNGQAPFTITPVFSDFPVH-----215

QY 229 VSGIOPLNEGVARCNESQGDVATCSQDCAASCPAIRPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESDEVTAPCSCDCSIVCGPKQPPPPAPWTLILGDAMYIM 272

QY 280 QMPGSLVLIILCSFVAVTILL-----VGRFVAPARDKSRMDPKGTSLS 326
Db 273 WITYMAFLVFFGAFFVWCYRKRYFVSYPIDSNIAFSV--NASDKG-----ASCC 324

QY 327 DKLSFSTHTLLGFOFGWGTWASWELTILVSVIPVVALAAGLVFTELTDPVLSWAP 386
Db 325 DPVSAAFEGCLRLFTWGSFCVRNCGVIFSLVFITACSGLVFVRVITTFVLDWSAP 384

QY 387 NSQARSEKAFHDQHGPFRTNQVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDQHGPFRTNQVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 444

QY 446 ELQERLRLHQVMSPEAQRNLSLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLTA 505
Db 445 DIQIALEN--ITASYDNETVTLQDICALPLSYNT---NCTLSLVNYPQNSHSDHKK 499

QY 506 NOTLMQOTSQVDWKDHFYCANAPLTFKDGTLALSCMADYGPVFPFLAIGSGYKDY 565
Db 500 GDDFF---VYADYHTEFLCYVRAPASLNDTSLHDPCLTGFGPFPVPLVGGYDDQYN 556

QY 566 EBAELIMTSLNYPAGDPRLLAQAKLWEAFLEEMRAFQRMFMQVFTTFAERSLEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAWKEBFINFKVKN---PNLTISFTAERSIEDEL 613

QY 626 NRTABDLPIFATSYIVIFLYISLALGSYSSRSRVVDSKATLGLGGVAVVLGAVWAMG 685
Db 614 NRESUSDVTVISVIAIMFYISLALGHYKSCRLLVDSKISGLIAGILVLSVACSLG 673

QY 686 FFSYLGIRSSVLQVVPFLVLSVADNIFI VLEYQRLPRPFGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTVIEVPELVAVGVDNIFILVQAYQRDERLQGETLQQLGRVLGEVAP 733

QY 746 SMLLCSLSBAICFFIGALTMPMPAVTRFALTSGVLVLDPLLOKSAFVALLSDSKRQAS 805
Db 745 SMLLCSLSBAICFFIGALTMPMPAVTRFALTSGVLVLDPLLOKSAFVALLSDSKRQAS 805
```

```

Db 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKQPPPPPPAPWITILGLDAMYIM 272
QY 280 QMPGSLVLIILCSVFAVVTILL-----VGRVAPARDKSKXVDPKKGTSLS 326
Db 273 WITYMAFLVFFGAFVAVCYKRYFVSVBTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSSTHTLLQFQCGWGTWASWPLTLVLVSVIPVVALAAGLVFTELTTDPVLSWAP 386
Db 325 DPVSAFEGCLRLRLETRMGSFCVRNPGCVIFPSLFTACSGLVFVRVTINPVDLWSP 384
QY 387 NSQARSEKAFDQHFQFPFRNQVILTAPNRSSRYDSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDQHFQFPFRTEQLIIRAPLTDKHIYQYPYSGADVFGPPLDIQILHQL 444
QY 446 ELQERLRHLQVMSPEAQRNLSQDICYAPLNPDNTSLYDCCINSLLQYQNNRTLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVLNYFQNSHVLHKK 499
QY 506 NOTLMGQTSQVMDKHFLYCANAPLTFKDGITALALSCMADYCAPVPPFLAIGYKGDYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLHDPCLGTGCGPVFPMVLVGGYDQNNY 556
QY 566 EAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLERAFQRRMAGMFQVTFTAERSIEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAKEKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYISLALGSYSWSRVMDSKATLGGVAVVLGAVMAAG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHKICSRLLVDSKVSLSGIAGILVLSSVACSLG 673
QY 686 PFSYILGRSSVLVQVWPFVLVSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVPLVAVGVDNIFILVQAYQDRERLQGETLDOQLGRVLGEVAP 733
QY 746 SMLLCSLSEAICFFLIGALTMPAVRTFALTSGSLAVILDFLLQMSAFVALLSLSKQRAS 805
Db 734 SMFLSSFTSEVATFFGALSVMVAHTFSLFAGLAVFIDFLQITCFVSLGLDIERQKEN 793
QY 806 RLDDVCCVKPQLPPPGQ-GEGLLAGFQKAYAPFLHWHITRGVLLFLFALFGVLSYM 864
Db 794 RLDFICCVRGADGTSVQASECLFRFPKNSVSLPLLDKWRMPVIAIFVGLSFI AVL 853
QY 865 CHISVGLDQELALPKDSYLLDFLNLRYFVAGPVYFVTLTLYNFSSEAGMNAICSSAG 924
Db 854 NKVDIGLQSLMPDSDYWDYFKSISQYLHAGPPVYFVLERGHDYTSKQGNVYCGGNG 913
QY 925 CNPFSFTQKIQVATEFPQSYLAIPASSWDDFDLWLT-SSCRLYISGPNKDKPCST 983
Db 914 CNNDSLVQOIFENAAQLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VDPACVR-CRLTPEGKORPGGDFMFLPMLSDNPNPKCGKGHAAYSAVNILLGH 1029
QY 1041 DQCVLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGTATFTWHTVLTQVTSADFDALKKARLIASNT-ETWINGS--AYRVFPYSVTV 1086
QY 1101 FVEQYLTILPEGLFMLSCLVPTTAVSCLLGLDRSLGNLLSITVILVDTVGFMAWLD 1160
Db 1087 FVEQYLTITDITFNLVGSIGAIFLVWTVLGCLEWSAVIMCATIAMVLVNMFGVWMLWG 1146
QY 1161 ISYNAVSLINLVSAGMSVEFVSHITRSPASTKTWLERAKETISMGSAVPAVAMTN 1220
Db 1147 ISLNAVSLNLMVSCGISVEFCSHITRAFTPMKGSRRVERAEALAHGSSVFSGITLK 1206
QY 1221 LPGAIVLGLAKAQLQIIFERLNLTLITLGLLHGLVFLPVILSVYVGPVWPA 1272
Db 1207 FGGIVVLFAKSIQIFIFRMYLANVLLGATHGLIFLVLUSYIGPSVKA 1258

```

RESULT 84

```

US-60-680-002-1038
; Sequence 1038, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1038
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-1038

Query Match 34.8%; Score 2402.5; DB 50; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLAQSEPYTHIQPGYCAYFDECG-----KNPELSGSLMTLSNVSCLS 59
Db 4 RGLALGLLILLCLCPAQVFSQ-----SCVMYGEGLAYGDKRYNCEVSG-----46
QY 60 NTPARKITGDHILILLOKI CPLYTGPTQOACCSAKOLVSEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELPCGFFFG-NVSLCCDVRQLQTLKDNLQPLQLFSLRSCPCFN 103
QY 120 FVNLHCHTCSNOSLFINVTR-----VAQLGAGQLPAVVAEAFQHSFAESYDSCSRV 175
Db 104 LNLFCETCSQROFLNVTATEDYVDPVTNQTNTKVLQYVYQGSFANAMYNACRDV 163
QY 176 RVPAAATLAVGTCMVGVSALCNAQRLNFQDGTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKDADA-CNATNWEYFWPKNGQAPPTITVPFSDFPVH-----215
QY 229 VGSQIPLNEGVARNESQGDVATPCSDCAASCAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSQDCSIVCGPKQPPPPPPAPWITILGLDAMYIM 272
QY 280 QMPGSLVLIILCSVFAVVTILL-----VGRVAPARDKSKXVDPKKGTSLS 326
Db 273 WITYMAFLVFFGAFVAVCYKRYFVSVBTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSSTHTLLQFQCGWGTWASWPLTLVLVSVIPVVALAAGLVFTELTTDPVLSWAP 386
Db 325 DPVSAFEGCLRLRLETRMGSFCVRNPGCVIFPSLFTACSGLVFVRVTINPVDLWSP 384
QY 387 NSQARSEKAFDQHFQFPFRNQVILTAPNRSSRYDSLLGPK-NFSGILDLLLELL 445
Db 385 SSQARLEKEYFDQHFQFPFRTEQLIIRAPLTDKHIYQYPYSGADVFGPPLDIQILHQL 444
QY 446 ELQERLRHLQVMSPEAQRNLSQDICYAPLNPDNTSLYDCCINSLLQYQNNRTLLTA 505
Db 445 DLQIAIEN--ITASVDNETVTLDICLAPLSPYNT---NCTILSVLNYFQNSHVLHKK 499
QY 506 NOTLMGQTSQVMDKHFLYCANAPLTFKDGITALALSCMADYCAPVPPFLAIGYKGDYS 565
Db 500 GDDFF---VYADYHFLYCVRAPASLNDTSLHDPCLGTGCGPVFPMVLVGGYDQNNY 556
QY 566 EAEALIMTFSLNNYPAGDPRLAQAKLWEAEFLERAFQRRMAGMFQVTFTAERSIEDEI 625
Db 557 NATALVITFPVNNYNDTEKQRAQAKEKEFINFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDLPFATSYIVIFLYISLALGSYSWSRVMDSKATLGGVAVVLGAVMAAG 685
Db 614 NRESDSVFTVVISVAINMFLYISLALGHKICSRLLVDSKVSLSGIAGILVLSSVACSLG 673
QY 686 PFSYILGRSSVLVQVWPFVLVSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVPLVAVGVDNIFILVQAYQDRERLQGETLDOQLGRVLGEVAP 733
QY 746 SMLLCSLSEAICFFLIGALTMPAVRTFALTSGSLAVILDFLLQMSAFVALLSLSKQRAS 805

```

```
Db 734 SMFLSSPSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDICKRQEN 793
Qy 806 RLDVCCVKPQELPPPGQ-GBGLLGGFPQKAYAPPELLHWITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFPKNSYSPLLLKDWMPRIVIAIFVGLVSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSAGHNAICSSAG 924
Db 854 NKVDIGLQSLSDMPDSDYMDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVCQGMG 913
Qy 925 CNMFSTQKIQVATEFPQOSYLAIPASSWVDDFIDWLTP--SSCRLYISGPNKDKFCPST 983
Db 914 CNNDSLVQOIFNAQAOLDNYTRIGFAPSSWIDDYFDMVAPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGAHAASAVNNILGH 1029
Qy 1041 DGOVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATFMTYHTVLTQTSADPIDALKKARLIASNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISVMILVDTVGFMAIWD 1160
Db 1087 FVEQVLTITDITFNLGVSLSGAIPLVTWVLLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRAEAEALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRNLNLITLGLLHGLVFLPVILSYVGPDPNPA 1272
Db 1207 FGGIVVLAPAKSQIPQIFFRMYLAMVLLGATHGLIFLPLVLSYIGPSVNKA 1258

RESULT 85
US-60-720-274-191
; Sequence 191, Application US/60720274
; GENERAL INFORMATION:
; APPLICANT: Olga Iakoubova
; TITLE OF INVENTION: Genetic Polymorphisms Associated with
; TITLE OF INVENTION: Cardiovascular Disorders and Drug Response, Methods of
; TITLE OF INVENTION: Detection and Uses Thereof
; FILE REFERENCE: CD000002
; CURRENT APPLICATION NUMBER: US/60/720,274
; CURRENT FILING DATE: 2005-09-23
; NUMBER OF SEQ ID NOS: 22965
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 191
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
US-60-720-274-191

Query Match 34.8%; Score 2402.5; DB 51; Length 1278;
Best Local Similarity 39.8%; Pred. No. 2.3e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGWLLWALLRLLAQSEPYTHIQPGYCAFYDECG-----KNPELSSGLMTLSNVSCLS 59
Db 4 RGLALGLLLLLCPAQFSQ-----SCVWYGEIGIAYGDKRYNCEYSG----- 46
Qy 60 NTPARKITGDHILTLQKICPRLYTGNTQACCSAKQOLVSLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYLDVQELCEGFFPG-NVSLCCDVRLQTLKDNQLPLQLSRCPSCFN 103
Qy 120 FVNLHCNTCSNQSLFINVTR-----VAOLGAGQLPAVVAYRAFYQHSFAEYSDSCSRV 175
Db 104 LLNLFCELTCSRQSQFLNVATATEDYVDPVTQTKNVKELQYVVGQSFPANAMYNACRDV 163
Qy 176 RVPAAATLAVGTCVGYGSAIACNAQRWLNPFQDGTGNGLAPLDIT-----PHLEPGQA 228
```

```
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEYMFNKONGQAPFTIITPVFSDFPVH----- 215
Qy 229 VSGSIQPLNEGVARCNESQDDVATCSCODCAASCAPAIARPQ-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGDSDVDEVTAPACSCODCSI VCGPKPQPPPPAPMTIILGLDAMYYIM 272
Qy 280 QMPCSLVLIILCSVEFAVVTILL-----VGFRVAPARDKSMVDPKKGTSL 326
Db 273 WITWMAFLLPFGAFFAVWCVRKRYFYSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSEFSTHTLIGQFQKGWTVASWPLTILVLSVPPVVALAAGLVFTELTTPDVELWSAP 386
Db 325 DPVSAAEGLCLRLFRWGSFCVENPGCVIIFSLVFITACSSGLVFVRVTINPVDLWSAP 384
Qy 387 NSQARSEKAFHDQHFQFPFRFTNQVILTAIPNRSSTRYDSLLGLPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEYFDQHFQFPFRTEQLIIRAPLTDKHIYQYPYSGADVPPGPPDLIQILHQVL 444
Qy 446 ELQERLRHLQVMSPEAQRNISLQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLLTA 505
Db 445 DLQIAIEN--ITASDYNETVILQICLAPLSFYNT---NCTILSVLNTFQNSHSLVDHKK 499
Qy 506 NOTLWGQTSQVDKWDHFLYCANAPLTFKDGFTALALSMDADYAPVFPFLAIGGYKGYKS 565
Db 500 GDDFF---VYADYHTHFLYCVRAPASLNDTSLLDHDPCLGTGPGVPFVWLVGGYDQNYN 556
Qy 566 EBAELIMTFSLNNTPAGDPRLAQAQKLWEAEFLKEMRAFORRMAGMFQVTFTAERSLEDEI 625
Db 557 NATALVITFPPNNYNTDEKLRQAQAEKEFINFVKYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPAPATSVIVILFVLSIALGYSWSRVMVDSKATLGLGGVAVVLGAVMAAG 685
Db 614 NRESQSDVFTVVISVAINFLYISIALGHIKCRLLVDSKVSGLIAGLILVLSVACSIG 673
Qy 686 PFSYGLIRSLVILQVVPFLVLSGADNIFIFVLEYQRLPRRPGPREVHIGRALGRVAP 745
Db 674 VFSYIGLPLTLIVIEVIFPLVLAGVDNI FILVQAYQDERLQGETLDOQLGRVLGEVAP 733
Qy 746 SMLCISLSEATCFPLGALTPMPAVTTPALTSGLAVIDFLLOMSAFVALLSDDSKROEAS 805
Db 734 SMFLSSPSETVAFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLDICKRQEN 793
Qy 806 RLDVCCVKPQELPPPGQ-GBGLLGGFPQKAYAPPELLHWITRGVLLFLFALFGVSLYSM 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFPKNSYSPLLLKDWMPRIVIAIFVGLVSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLGYNFSSAGHNAICSSAG 924
Db 854 NKVDIGLQSLSDMPDSDYMDYFKSISQYLHAGPPVYFVLEBEGHDYTSKQGNMVCQGMG 913
Qy 925 CNMFSTQKIQVATEFPQOSYLAIPASSWVDDFIDWLTP--SSCRLYISGPNKDKFCPST 983
Db 914 CNNDSLVQOIFNAQAOLDNYTRIGFAPSSWIDDYFDMVAPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSRVPSVEQFHKYLPWFLNDRNPKCPKGLAAYSTSVN--LTS 1040
Db 971 VVDPACVR-CRPLTPEGKQRPQGGDFMRFLPMFLSDNPNPKCGKGAHAASAVNNILGH 1029
Qy 1041 DGOVLASRFMAYHKPKNSQDYTEALRAARELANITADLRKVPCTDPAFEVFPYITNV 1100
Db 1030 GTRVGATFMTYHTVLTQTSADPIDALKKARLIASNT-ETWINGS--AYRVFPYSVFYV 1086
Qy 1101 FVEQVLTILPEGLFMLSCLVPTFAVSCLLGLDLRLSGLLNLISVMILVDTVGFMAIWD 1160
Db 1087 FVEQVLTITDITFNLGVSLSGAIPLVTWVLLGCELWSAVIMCATIAMVLVNMFGVWMLWG 1146
Qy 1161 ISYNAVSLINLSVAGMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLVNLVMSCGISVEFCSHITRAFTVSMKGSRAEAEALAHMGSSVFSGITLTK 1206
Qy 1221 LFGILVLGLAKAQLIQIFFRNLNLITLGLLHGLVFLPVILSYVGPDPNPA 1272
```

```
Db 1207 FGGIVLAFPAKSIQIFQIFPRMVLAMVLLGATHGLIFLPVLLSYIGPSVNA 1258

RESULT 86
US-11-222-021-440
; Sequence 440, Application US/11222021
; GENERAL INFORMATION:
; APPLICANT: Bruno DOMON et al.
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001548
; CURRENT APPLICATION NUMBER: US/11/222,021
; CURRENT FILING DATE: 2005-09-09
; NUMBER OF SEQ ID NOS: 7554
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-222-021-440

Query Match 34.8%; Score 2402.5; DB 42; Length 1279;
Best Local Similarity 39.8%; Pred. No. 2.4e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLLAQSBPYTHIQPGYCAFYDECG-----KNPELSSGLMTLSNVCSLS 59
Db 4 RGLALGLLLLLCPAQVFSQ-----SCWYGEIGYADKRYNCEYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQIVLSLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLLCCDVRLQTLKONLQLPLQFLSRCPSCFYN 103

QY 120 FVNLHCHNTCSNQSFLINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPQSQFLNVTATEDYDVPNTQTKNVKELQYVYGQSFAANMYNACRDV 163

QY 176 RVPAANTAVGTWCGYVGSALCNAORLNFQSDTGNGLAPLDIT-----FHLLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNWIETMFKNDGQAPFTITPVFSDFPVH-----215

QY 229 VSGGIQPLNEGVARCNESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 --GMEPMNATKGCDESDEVTPACSCQDCSIVCGKFPQPPPPAPWTILGLDAMYIM 272

QY 280 QMPGSLVLIILCSVFAVVTILL-----VGRVAPARDKSMVDPKGTLSLS 326
Db 273 WITYMAFLVFFGAFVAVWCYRKYRVSEYTPIDSNIAFSV--NASDGE-----ASCC 324

QY 327 DKLSFSTHTLLGQFGQGWGTWASWELTILVLSVIPVVALAAGLVFTLTDPVLEWSAP 386
Db 325 DPVSAAFEGCLRLFRMGSCVRNFGCVIFFEFLVITACSSGLVFRVTTNPVDLWSAP 384

QY 387 NSQARSEKAFDHQFGPFRFTNQVIITAPNRSRYVDSLLGPK--NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDQHFQGFRTTEQLIIRAPLTDKHIYPYPSGADVPFGPDLQIILQVL 444

QY 446 ELQERLRHLQVMSPEARNLSDICVAPLNPNTSLYDCCINSLLQYQNNRTLLTLTA 505
Db 445 DIQIALEN--ITASYNDETVTQDICIAPLSPYNT---NCTILSVNLVFNQNSHVLDDHK 499

QY 506 NOTLMQTSQVMDKHFLYCANAPLTFKDGITALALSCMADYAGAPVFPFLAIGYKGOYS 565
Db 500 GDDFF---VYADYHTFLYCVRAPASINDTSLHDFCLGTGFGVPFWLVGLGYDDQNTN 556

QY 566 EBAELIMTSLNYPAGDPRLOAKLWEBAFLEEMPAFORRMAGMFQVTTAERSLEDEI 625
Db 557 NATALVITPVNNYNDTEKLQRAQAWKEKFINFVNKYN---PNLTISFTAERSIEDEL 613

QY 626 NRTTAEDLPATSYIVIFLYTSLALGSSYSRSMVDSKATILGCGVAVLGAVMAMG 685
Db 614 NRESDDVFTVVISYAIMFLYISLALGHITKSCRRLLVDSKVISGIAGILIVLSSVACSLG 673

QY 686 FFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLEYORLPRRPGEPREVHIGRALGRVAP 745

Db 674 VFSYIGLPLTLIVIEVIFLVAAGVDNIFILVQAYQRDERLQGETILDQQLGRVLGVAP 733
QY 746 SMLLCSLSRAICPFLGALTTPMPAVRTFALTSGLAVIDLFLLOMSAFVALLSLSKQORAS 805
Db 734 SMLSSFSFSETVAFPLGALSVMPPAVHTFSLFAGLAVFIDFLQLITCFVSLILGLDIKROKN 793
QY 806 RLDVCCCVKPELPPPGQ--GEGLLGFFQKAYAPFLLHMITRQGVLLLLFLALFGVSLYSYM 864
Db 794 RLIDFCCVRGABDGTSVQASESCLPFRFKNSYSPLLKKOMMRPIVIAIFGVLSFSJAVL 853
QY 865 CHISVGLDQELALPKDSYLLDYFLFNRYFEVGAIVYFVTTLYGNFSSESAGMAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMYDYFKSISQYLHAGPPVYVLEEGHDYTSRSGQNMVCGMG 913
QY 925 CANNFSPTOKIYATEFPEQSILAIAPASSWDDDFIDMLTP--SSCCRLYISGPNKDKPCPT 983
Db 914 CNDLSLVQQLFNAAQLDNTRIGFAPSSWIDDFDMVKPOSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLCKNCMSIT--MGSVRPSVEQPHKYLPMFLNDRPNIKCPKGLLAAYSTSVN--LTS 1040
Db 971 VVDPAVCR--CRPLTPGKQRPQGGDFWRFLPMLSDNPNPKCGKGHAAVSSAVNILLGH 1029
QY 1041 DGOVLASRFMAYHKPLKNSQDYTEALRAARELANITADLRKVPGTDPAPFVFPYTTNV 1100
Db 1030 GTRVGATYPMYHTVLTQTSADFDALKARLIASNTV--ETMGINGS--AVRVFPYSVFYV 1086
QY 1101 FRYOYLTIIPGFLPMLSLCLVPTFAVSCILLGLDLSGLLNLLSIVMILVDTVGFMALWD 1160
Db 1087 FRYOYLTIIDDTIFNLGSLGAIFLVTWLLGCELWSAIVMCATIAMVLNFMGVMMWLMG 1146
QY 1161 ISYNAVSLNLVSAVQMSVEFVSHITRSPAISTKPTWLERAKEATTSMGSAVAGVAMTN 1220
Db 1147 ISLNAVSLNVLWMSGIVSEFCHITRAFTVSKGSRVERAEALAHMGSVSFSGITLTK 1206
QY 1221 LPEGLVLGLAKAQLIQIFPFRNLMLTLTLGLLHGLVFLPVLVSYVSGPDVNP 1272
Db 1207 FGGIVLAFPAKSIQIFQIFPRMVLAMVLLGATHGLIFLPVLLSYIGPSVNA 1258

RESULT 87
US-60-680-002-1037
; Sequence 1037, Application US/60680002
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001601
; CURRENT APPLICATION NUMBER: US/60/680,002
; CURRENT FILING DATE: 2005-05-12
; NUMBER OF SEQ ID NOS: 2417
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1037
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-680-002-1037

Query Match 34.8%; Score 2402.5; DB 50; Length 1279;
Best Local Similarity 39.8%; Pred. No. 2.4e-201;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

QY 7 RGWLLWALLRLLAQSBPYTHIQPGYCAFYDECG-----KNPELSSGLMTLSNVCSLS 59
Db 4 RGLALGLLLLLCPAQVFSQ-----SCWYGEIGYADKRYNCEYSG-----46

QY 60 NTPARKITGDHLLILLOKICPRLYTGNTQACCSAKQIVLSLEASLSITKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG--NVSLLCCDVRLQTLKONLQLPLQFLSRCPSCFYN 103

QY 120 FVNLHCHNTCSNQSFLINVTR----VAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRV 175
Db 104 LNLFLCELTCSPQSQFLNVTATEDYDVPNTQTKNVKELQYVYGQSFAANMYNACRDV 163

QY 176 RVPAANTAVGTWCGYVGSALCNAORLNFQSDTGNGLAPLDIT-----FHLLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA--CNATNWIETMFKNDGQAPFTITPVFSDFPVH-----215

QY 229 VSGGIQPLNEGVARCNESQDDVATCSQDCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 --GMEPMNATKGCDESDEVTPACSCQDCSIVCGKFPQPPPPAPWTILGLDAMYIM 272

QY 280 QMPGSLVLIILCSVFAVVTILL-----VGRVAPARDKSMVDPKGTLSLS 326
Db 273 WITYMAFLVFFGAFVAVWCYRKYRVSEYTPIDSNIAFSV--NASDGE-----ASCC 324

QY 327 DKLSFSTHTLLGQFGQGWGTWASWELTILVLSVIPVVALAAGLVFTLTDPVLEWSAP 386
Db 325 DPVSAAFEGCLRLFRMGSCVRNFGCVIFFEFLVITACSSGLVFRVTTNPVDLWSAP 384

QY 387 NSQARSEKAFDHQFGPFRFTNQVIITAPNRSRYVDSLLGPK--NFSGILDLDLLELL 445
Db 385 SSQARLEKEYFDQHFQGFRTTEQLIIRAPLTDKHIYPYPSGADVPFGPDLQIILQVL 444

QY 446 ELQERLRHLQVMSPEARNLSDICVAPLNPNTSLYDCCINSLLQYQNNRTLLTLTA 505
Db 445 DIQIALEN--ITASYNDETVTQDICIAPLSPYNT---NCTILSVNLVFNQNSHVLDDHK 499

QY 506 NOTLMQTSQVMDKHFLYCANAPLTFKDGITALALSCMADYAGAPVFPFLAIGYKGOYS 565
Db 500 GDDFF---VYADYHTFLYCVRAPASINDTSLHDFCLGTGFGVPFWLVGLGYDDQNTN 556

QY 566 EBAELIMTSLNYPAGDPRLOAKLWEBAFLEEMPAFORRMAGMFQVTTAERSLEDEI 625
Db 557 NATALVITPVNNYNDTEKLQRAQAWKEKFINFVNKYN---PNLTISFTAERSIEDEL 613

QY 626 NRTTAEDLPATSYIVIFLYTSLALGSSYSRSMVDSKATILGCGVAVLGAVMAMG 685
Db 614 NRESDDVFTVVISYAIMFLYISLALGHITKSCRRLLVDSKVISGIAGILIVLSSVACSLG 673

QY 686 FFSYLGIRSLVLQVVPFLVLSVGADNIFIFVLEYORLPRRPGEPREVHIGRALGRVAP 745
```



```

Db 602 ALVITFPVNNYNDTERLQRAWKEFISFVKYKN---PNTISFTAESRSEDENRE 658
QY 629 TAEDLPFATSYIVIFLYISLGSYSRSMVMDSKATGLGVAVVLGVAAMAGPFS 698
Db 659 SNSDVFVTVIISYVVMFLYISLALGHISQCSRLVDKISLGIAGILVLSVACSLGIFS 718
QY 689 YLGIRSLVLOVVPFLVSVGADNIFIFVLEYORLPRRPGEPREHIGALGRVAPSM 748
Db 719 YMGPMPLTVIVIEVIFLVAVGVDNIFILVQYORDERQEEETDQOLGRILEVAPTWF 778
QY 749 LCSLSAICFFLALTPMPAVRTPALTSGLAVILDLLQMSAFVALLSLDSKROEASRLD 808
Db 779 LSFSETSAPFFGALGSMFAVHTFSLFAGMAVLIDFLQLITCVLSLGLDIKRQEKHLD 838
QY 809 VCCCVKQELPPPGQG----EGLLGFQKAYAPFLHMITRIGVILLFLALFGVLSYM 864
Db 839 ILCVRGAD---DGQSHASESYLFRFFKNYPAPLILLKOWLRPIVVAVFVGLSPSAVV 895
QY 865 CHISVGLDQELALPKDSYLLDYFLNRPFEVGAPEYFVTTLYNFSSEAGMAICSSAG 924
Db 896 NKVDIGLQSLMPNDSYVIANFKLAQYLHSGPPYFVLEEGYNSRRKGQNMVCGMG 955
QY 925 CNNFSFTQIOYATEPEQSYLAIPASSWVDDFDIDLTP--SSCCRLYISGNPKDKPCPST 983
Db 956 CNDLSLVQILFNAEALDITTRVGFAFSSWIDIDFDWVSQSSCCRLY---NVTHQFCNAS 1012
QY 984 VNSLNLKNCMSIT-MGSVRPSVEQFKYLPWFINDRPNIKCPKGLAAYSTSVNLTSQG 1042
Db 1013 VMDPTCVR-CRPLTPGKQRPQCKEFMKFLPMLSDNPNPKCGKGHAAVGSANVIGDD 1071
QY 1043 QVL-ASRPMAYKPLKNSODYTALRAARELANITADLRKVGTDPAPFVPPYITNVF 1101
Db 1072 TVIGATYFTYTHILKTSADYTDAMKARLIASNITETWRS-KGSD--YRVPEYSFVYF 1128
QY 1102 YEOYLILPEGLPMLSLVPTFAVSCLLGLDLRLSGLNLILSVMIIVDTVGFMAIWDI 1161
Db 1129 YEOYLITIDITFNLVSLSGSIPLVTLVLGCLWSAVIMCITIAMILVNMFGVMWLGI 1188
QY 1162 SYNANSLINLVAGMSVFSVSHITRSFAISTKPTWLERAKEATISMGSAVAGAVMTNL 1221
Db 1189 SLNANSLVNLVMSGISVFCSHITRAFTWSTKGSRSRABEALAHKMSVSGITLTKF 1248
QY 1222 PGLVLGLAKAQIQTFFRLNLLITLGLLHGLVLPVLVSYVGPDPNPA 1272
Db 1249 GGVIVLAFAKSQIFEIFRMYLAMVLLGATHGLILFLPVLLSYIGPSVKA 1299

RESULT 89
PCT-US03-24982A-291
; Sequence 291, Application PC/TUS0324982A
; GENERAL INFORMATION:
; APPLICANT: Syngenta Participations AG
; APPLICANT: Stam, Lynn
; APPLICANT: Kamdar, Kim
; APPLICANT: Spana, Eric
; APPLICANT: Bachmann, Jane
; TITLE OF INVENTION: Nucleic Acid Sequences from Drosophila Melanogaster that Encode F
; FILE OF INVENTION: Essential for Viability and Uses Thereof
; FILE REFERENCE: 70131WOPCT
; CURRENT APPLICATION NUMBER: PCT/US03/24982A
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: 60/422,377
; PRIOR FILING DATE: 2002-10-29
; NUMBER OF SEQ ID NOS: 381
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 291
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
PCT-US03-24982A-291

```

Query Match 32.7%; Score 2262; DB 1; Length 1287;
 Best Local Similarity 37.8%; Pred. No. 6.1e-189;

```

Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;
QY 10 LLWALLRLRAQSEPYTTIHQPGYCAFYDECGKPELSGLMTLSNVSCLSNTPARKITGD 69
Db 17 LLAAVLPTLIQSSKOD-----CWYGVNTN-DESHS-----QNCYPNGTAKEMAYD 62
QY 70 HLILLOKICPLRYGTGNTQACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTC 129
Db 63 GLELLKRCGFLLENSEKFCDDKNQVELLNKNVLAGNILDRCPSCMENLVRIHQFTC 122
QY 130 SPNOSLFINVTRVAQLGAGLPVAVYEAIFYQHSFAEQSDYSCSRVRVPAATAVLAVGTM 189
Db 123 SPQASFMHVAVTQNNKGD-EYISSVDLHISTEYINKTYKSCQSVFQGTQGLAFDLMC 181
QY 190 GYVSALCNAQRLWLFQGTGNGLAPLDTTFHLLRPGQAVGSGIQPLNEGVARNESQGD 249
Db 182 GAYSASRCNPTKWFNMGDATWYVPFQIYIQHEP-KSNSNFTPLANTVTPCQAVSS 240
QY 250 DVATSCQDCAASC-----PAIARPO-----ALDSTFYLQMPGSLVLIILCSVFAV-VT 299
Db 241 KLPACSCSDCLSCQGPPEPPRPPPKIVGLDAVF-----VIMAAVFLVGLV 288
QY 300 ILLVG---PRVAPARDKSKMVDPKGTSLSDKLSFS-----THTLLGQFQGW 344
Db 289 VFLMSGFLFTQSSMDDNFQVD---GNVSDEMPYSENDSYFEKLGHAHTTETLETFTKW 345
QY 345 GTWVASWPLTILVLSVIPVVALAAGLVFTELTDPVELWSAPNSQARSKAHDFHQPF 404
Db 346 GTFYASNPGTLIAGASLVILGYGINFIEITDPVKLWASPNKSLREREFDTFSPF 405
QY 405 FRTNQVILTPAPNRSSRYRVSLLGPKNFSGIILDLLELLELQERLRLHQMSPBAQRN 464
Db 406 YRLEQIIIIKAVNLPOIVHNT-SNGPYTFGPVDRBFLTKVLDLQEGIKEINA-----NG 458
QY 465 ISLDICYAPLNDPNTSL--YDCCINSLLQYQNNRTLLLLFANQTLMGQTSQVDWDXF 522
Db 459 TQLDKICYAPLSDGSEIDVSCVQVQSWGFGDDRERL-----DDHDENGENVTVYLDAL 514
QY 523 LYCANAPLTFKGTALALSCMADYCAPVFPFLAIGY-----KGKDYSAEALIMTF 574
Db 515 YDCISNPLY-----CLAPYGPVDPALALGGLFPGLPQDLTGSTKPELANAILTF 564
QY 575 SLNNYPAGDPRLAQAKLEAEFLEENRAF-QRRMAGMFQVTFTAERSLDEINRTTAEDL 633
Db 565 LVKNH-HNKTDLLENALTWEKKVFEMTNTYKNMSQYMDIAFTSERSEIDELNRSSQSDV 623
QY 634 PIFATSYIVIFLYISLALGSYSRSMVMDSKATGLGVAVVLGVAAMAGPFSYLGTR 693
Db 624 LTLVSYLIMFYIAISLGHVKEPKRVFIDSKITLIGGVIVLASVWSVGVFGYGLP 683
QY 694 SSLVILQVVPFLVLSVGDNIIFIVLEYORLPRRPGEPREHIGALGRVAPSMILCSLS 753
Db 684 ATLIIIVEIPFLVAVGVDNIFILVQTHORDORKNETLEQQVGRILGKVGSMILTSLS 743
QY 754 EACFFLGLALTPMPAVRTPALTSGLAVILDLLQMSAFVALLSLDSKROEASRLDVCCCV 813
Db 744 ESFCFFLGLSLDMPAVRALVAGVALIIDFLQITCVLSLFTLTKREERNMDICCFI 803
QY 814 KPQELPPPGQGGELLGFFQKAYAPFLHMITRIGVILLFLALFGVLSVMCHISVGLDQ 873
Db 804 KGKPDSTITSNEGLLYKFPSSVYVVPFLMKKIVRASVMVFFAWLCFSAIAPRIDGLDQ 863
QY 874 ELALPKDSYLLDYFLNRPFEVGAPEYFVTTLYNFSSEAGMAICSSAGNNFSFTQK 933
Db 864 ELAMPQDSFVLHYFQSLNENLIMGPVYFVLKGLDAYTNSSDNLVACQYCNDSVLQY 923
QY 934 IOYATEFPEQSYLAIPASSWVDDFDIW-LTPSSCCRLYISGNPKDK--PCPSTVNSLNC 990
Db 924 IYLAERHSNQTIIARPASSWIDIDYFDWAAAASCCYK-----RKSGDFCPH--QDTSCL 976
QY 991 KNCMSITMGS-VRPSVQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVLI---A 1046
Db 977 R--CNITKNSLLRPEKEFVKYLPFLKXNDPDDTCAKAGHAAVGGAVRYSNSHERLNIEA 1034

```

```
QY 1047 SRFMAYKPLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAREVEPYTITN 1099
Db 1035 SYFMAYHTLKSSADYFLAESARKISANITQMLQRLMSNGVPMASALTVEVFPYSVFY 1094
QY 1100 VFEYQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLMLLSIVMLVDTVGFMAW 1159
Db 1095 VFEYQYLTMWSDTLQSMGIVLSIFVTVFLMGFDVHSALVVVITITMIVNLGGLMYW 1154
QY 1160 DISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAVAMT 1219
Db 1155 NISLNAVSLVNLVMAVGISVEFCSHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214
QY 1220 NLPGLVNLGLAKAQLIOTFFRLNLLITLLGLLHGLVFLPVTLISVVGVDVNPALAEQKR 1279
Db 1215 KPAGILVLFAKSQLFOVFYFRMYLGI VVIGAAHGLIFLPLVLLSYIGAPVSNARLRYHSQ 1274
QY 1280 A-----EEAFAAM 1288
Db 1275 AAHEHETALAGIL 1287

RESULT 90
US-09-614-150-12003
; Sequence 12003, Application US/09614150
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150
; CURRENT FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150-12003

Query Match 32.7%; Score 2262; DB 26; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;
QY 10 LLMALLRLAQSEPYTHIQPGVCAPYDECGKNPELSGLMILSNVSLNTPARKITGD 69
Db 17 LIAAVLFTLIQSSKQD-----CWVYGVCTNN-DPSSH-----QNCYPNGTAKEMATD 62
QY 70 HLILKQICPRLYTGNTQACSAKOLVSEASLSITKALLTRCPACSNFVNLHCHNTC 129
Db 63 GLELLKKRCGFLLENSKFPCCDKQVLLNKNVLAGNLDRCPCSMENLVRIHCQFTC 122
QY 130 SPNQSLFINVTRVAQIAGAGLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATAVAGTWC 189
Db 123 SPKQAEFMHVATQKNKGD-EYISSVDLHISTEYINKYKSCSQSVSPQTGQLADLMC 181
```

```
QY 190 GVTYSALCNAORWLNFGQDTGNGIAPLDITPHLEPGQAVGSGIOPNLNGVARCNESQSD 249
Db 182 GAYGASRCNPTKWNFGDATNPVPOITYIQHEP-KSNNNFTPLNVTTPCNOAVSS 240
QY 250 DVATCSCODCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIIILCSFAV-VT 299
Db 241 KLPACSCSDCLSCFPQGPPEPPPEPKIVGLDAYF-----VIMAAVFLVGL 288
QY 300 ILLVG----FRVAPARDKSKMVDPKKSGISLSKLSFS-----THILLGQFFQW 344
Db 289 VFLMGSLFTLQSGSMDNDNFQVD--GNDVSDEMPYSENDSYFEKLGATHTFLETFFTKW 345
QY 345 GTWASWPLTTLVLSVPVVALAAGLVPTLTPDVELWSAPNSOARSEKAFDHOHGF 404
Db 346 GTTFASNGLTLIAGASLVVILGYGINFIEITTPVKUWASPNSKRLEREFDFTKFEP 405
QY 405 FRTNQVILTAPNRSSRYDSILLGPKNPSGILDLLLELLELQBLRHLQWSPEAQRN 464
Db 406 YRLEQIIIKAVNLPIQIVHNT-SNGPYTPGVPDFREFLTKVLDLQSGIKEINA-----NG 458
QY 465 ISLQICVAPLNPDNTSL--YDCCINSLQYFQNNRTLLLTANQTLMGQTSQVNDKHF 522
Db 459 TQLKDCICVAPLSDGSEIDVSCVQSVGWYFGDDRERL-----DDHDEDNGFNVTYLDAL 514
QY 523 LYCANAPLTFKGTALALSCWADYGAPVPFLAIGGY-----KGKDYSEAEALIMTF 574
Db 515 YDCISNPYL-----CLAPYGGVPDPAIAGGLFPGDQLTGSTKPELANAIIITF 564
QY 575 SLNNYPAGDPRLAQAKLWEEAPLEBEMRAP-QRRMAGMPQVTTFAERSLEDEINRTABDL 633
Db 565 LVKNH-HNKTDLLENALTWEKKFVEMTNYTKNNMSQYMDIAFTSERSIEDELNRESQSDV 623
QY 634 PIPATSYIVIFYISLALGSYSSWSRVWDSKATILGLGVAVVLGAVMAAMGFYLGIR 693
Db 624 LTIIVSYLIMFYIAISLGHVKFVKRVIDSKITIGIGVIVLASVSSVGVFYIGUP 683
QY 694 SSLVTLQVVPVLVLSVGADNIPFVLEYQRLPRRPGEPREVIHIGALGRVAPSMLLCSLS 753
Db 684 ATLIIIVEIPFLVLAVGVDNIPILVQTHQRORKEVETLEQQVGRILKGVSMLLTSL 743
QY 754 EACFFLGALTPMPAVRTPALTSLGLVILDFLLQMSAFVALLSLDSKQEAERLDVCCV 813
Db 744 ESPCFFLGSLDMPAVRAFALYAGVALLIIDFLQITCFVSLFTLDTKREENRMDICCFI 803
QY 814 KPQELPPPGQGBGLLGFQKAYAPFLHWTIRGVVLLFLALFGVLSYMSCHISVGLDQ 873
Db 804 KGKKPDSITSNEGLLYKFFSSVYVFLMKIVRASVMVIFFAWLCPSTAIAPRIDGLDQ 863
QY 874 ELALPKDSYLLDYFLFNARYPEVGAPVYFVTLTYNFSSEAGMNAICSSAGCNNSFTOK 933
Db 864 ELAMPQDSFVLHYFOSLNNENLIGPVTVFLKGDLAYTNSSQNLVLCAGQYCNDSVLQ 923
QY 934 IQYATEFPPEQSYLAIPASSWDDFDIW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNC 990
Db 924 IYLASRHSNQTYIARPASSWIDDDYFDMAAAASCCCKY-----RKDSGDCPH--QDTSCL 976
QY 991 KNCMISITMGS-VRPSVEQFHKLTPWFLNDRPNKICPKGLAAYSTSVNLTSGQVL--A 1046
Db 977 R--CNITKNSLLRPEKEFVKYLPFLKDNPNDDTCAKAGHAAYGAVRYSNSHERLNIEA 1034
QY 1047 SRFMAYKPLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAREVEPYTITN 1099
Db 1035 SYFMAYHTLKSSADYFLAESARKISANITQMLQRLMSNGVPMASALTVEVFPYSVFY 1094
QY 1100 VFEYQYLTLPEGLFMLSCLVPTFAVSCLLGLDLRLSGLMLLSIVMLVDTVGFMAW 1159
Db 1095 VFEYQYLTMWSDTLQSMGIVLSIFVTVFLMGFDVHSALVVVITITMIVNLGGLMYW 1154
QY 1160 DISYNAVSLINLVSAVMSVEFVSHITRSFAISTKPTWLERAKEATISMGSAVFAVAMT 1219
Db 1155 NISLNAVSLVNLVMAVGISVEFCSHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214
```

QY 1220 NLPGLVGLAKAQLIQIFPRNLNLTLLGLLHGLVFLPVLISYVGPDPNPALEQR 1279
DB 1215 KPAIGLVAPAKSQIPQVYFRMYLGIIVVGAHGLIFLPLVLLSYGAPVSNARLYHSQ 1274
QY 1280 A-----EEAVALM 1288
DB 1275 AAHEHETALAGIL 1287

RESULT 91
US-09-614-150A-12003
; Sequence 12003, Application US/09614150A
; GENERAL INFORMATION:
; APPLICANT: venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150A
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150A-12003

Query Match 32.7%; Score 2262; DB 26; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LIAVLRLAQSEPYTTIHQPGYCAFYDECGNPELGSMTLSNVCLSNTPARKITGD 69
DB 17 LIAVLRLAQSEPYTTIHQPGYCAFYDECGNPELGSMTLSNVCLSNTPARKITGD 62
QY 70 HILLOKICPRLYTGNTQACSAKQLVSLASLITKALLTRCPACSNFNLHCHNTC 129
DB 63 GLELLKRCGFLLESENKFCCKQKQVLLNKNLGNLILDRCPSCMENLVRIHQFTC 122
QY 130 SPQSLFINVTRVAQLGACQLPAVVAEAFYQHSFAEQSYDSCSRVRVPAATLAVGTCM 189
DB 123 SPQSLFINVTRVAQLGACQLPAVVAEAFYQHSFAEQSYDSCSRVRVPAATLAVGTCM 181
QY 190 GYVGSALCAORLWLPQGTGNGLAPLDITPHLEBPQAVGSGIOPLNEGVARNCSQGD 249
DB 182 GAYSASRCNPTKWFNMGDATHPYYPQIYIQHEP-KSNSNFTPLNVTTPCNQAVSS 240
QY 250 DVATSCDCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSFVAV-VT. 299
DB 241 KLPACSCDCLSCQPGPEPRPPEPKIVGLDAYF-----VIMAAVELVGL 288
QY 300 ILLVG---PRVAPARDKSNVDPKGTSLSDKLSFS-----THLLGQFFQGM 344
DB 289 VFLMGSLFTQSSMDNFDQVD---GNDVSDMPYSENDSYFEKLGHAHTETFTFTFTK 345

RESULT 92

US-11-097-143-12003

; Sequence 12003, Application US/11097143

; GENERAL INFORMATION:

QY 345 GTWVASWELTILVLVSVIPVLAAGLVFTTELTTDPVELMWSANPSQARSSEKAFHDHQFGB 404
DB 346 GTYFASNPGCLTLIAGASLVILGYGINFETITTDVFKLWASPNKSRLEEFEDTKFSPF 405
QY 405 PRFTNOVILTAPNRSSRYRVDLSLLGPKNFGIILDLLELLELQERLRLHQWSPQAORN 464
DB 406 YLEQIIIIKAVNLPOIVHT--SNGPYTFGPVDFRFLTKVLQLOQEGIKINA-----NG 458
QY 465 ISLODICYAPLNPNTSL--YDCCINSLLQYQNNRTLLLTANOTLMCQTSOVOWKDH 522
DB 459 TQLKIDICYAPLSDGSEIDVSCVQSIWGYGDDRERL-----DDHEDNGFNVTYLDAL 514
QY 523 LYCANAPUTFFKDTALALSCMADYGAPVPPFLAIGY-----KGKDYSAEALIMTF 574
DB 515 YDCISNPLY-----CLAPYGGPVDPAIAGLFGFLPPGDLTGSTKFELANAILTF 564
QY 575 SLANNYPAGDPRLAQAKLWEEAELEEMRAP--QREMGMFQVTTAERSLEDEINRTAEDL 633
DB 565 LVKNH--HNKTDLENALTWEKPFVEFTNTYTKNMSQYMDIAFTSERSIEDELNRBSQSDV 623
QY 634 PIFATSYIVIFLYISIALGSSYSSMRVMDSKATLGLGVAVVLGAVMAAMGFFSYLIGIR 693
DB 624 LTILVSYLIMFYIAISLGHVKEFKRVFIDSKITLIGGVIIIVLASVSSVGVGIGLP 683
QY 694 SSLVILQVPLVLSVGADNIPFVLEYQRLPRRPGEPREVNHIGRALGRVAPSMILCSLS 753
DB 684 ATLIIIVEIPFLVLAAGVDNIFILVQTHORDQKNETLEQQVGRILKGVPSMLLTSL 743
QY 754 EACIPELGNLTMPAVRTEALTSGLAVILDFELQMSAFVALLSLDSKROEASRLDVCCCV 813
DB 744 ESFCFFLGLSDMPAVRAPALYAGVALLIDFLQITCFVSLFTLDTKRRENRMDCCPI 803
QY 814 KPQELPPPQCGEGLLGPFQKAYAPFLHWTGRGVLLFLALFGVLSYSMCHIISVGLQ 873
DB 804 KGKXPDSTISNEGLLYKFSVVYVFLMKKIVRASVMVIFPAWLCFSAIAIRIDIGLQ 863
QY 874 ELALPKDSYLLDYFLFLNARYFEVGAPVYVTTLYGNFSSEAGNAICSSAGCNFSFTOK 933
DB 864 ELAMPQDSFVLHYFQSLNENLNIIGPVYFVLKGLDAYTNSSDQNLVCAQYCNDDSVLTQ 923
QY 934 IQYATEPPEQSYLAIPASSWDDFDW--LTSSSCCLYISGNKDK--FCPSTVNSLCL 990
DB 924 IYLSRHSNQTYIARPASWIDDDYDMAAASSCKY-----RKDSGDFCPH--QDTSCL 976
QY 991 KNCMSITWGS--VRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQDQVL---A 1046
DB 977 R--CNITKNSLLRPEKEFVKYLPFLKONPDCTAKAGHAAYGAVRYSNSHERLNIEA 1034
QY 1047 SRFMAYKPLKNSQDYTEALRAARELANITADLR-----KVP--GTDPAFVFPYTTIN 1099
DB 1035 SYFMAYHTILKSSADYFLALBSARKISANITQMLQGLMSNGVPMASALTVEVPVSVFY 1094
QY 1100 VPEYQYLILPEGLPMLSLCLVPTTAVSCLLGLDRSGLLNLISIVMLLVDVTGFMALW 1159
DB 1095 VPEYQYLTMWSDTLQSMGISVLSIFVVTFLMGPDVHSALVVVITITIVVNLGGLMYM 1154
QY 1160 DISYNAVSLINIVSAGMSVBFVSHITRSPALSTRPTWLERAKBATISGSAFVAGVAMT 1219
DB 1155 NISLNAVSLVNLVMAVGSVEFCSHLVSFATSKVSQIDRAANDSLSKMGSIFSGITUT 1214
QY 1220 NLPGLVGLAKAQLIQIFPRNLNLTLLGLLHGLVFLPVLISYVGPDPNPALEQR 1279
DB 1215 KPAIGLVAPAKSQIPQVYFRMYLGIIVVGAHGLIFLPLVLLSYGAPVSNARLYHSQ 1274
QY 1280 A-----EEAVALM 1288
DB 1275 AAHEHETALAGIL 1287

```

; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/170,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12003
; LENGTH: 1287
; TYPE: PRF
; ORGANISM: DROSOPHILA
US-11-097-143-12003

Query Match          32.7%; Score 2262; DB 40; Length 1287;
Best Local Similarity 37.8%; Pred. No. 6.1e-189;
Matches 504; Conservative 244; Mismatches 469; Indels 116; Gaps 30;

QY 10 LLWALLRLAQSEPTTHIQGYCAFYDECGNPNLSGLMTLSNVSCLSNTPARKITGD 69
DB 17 LIAVLFTLIQSKQD-----CWVYGVCTN-DFSHS-----QNCFYNGTAKEMATD 62
QY 70 HUILOKICPRLYTGENTQACSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTC 129
DB 63 GLELLKRCGFLLESENKFCCKQKQVELLNKNVLAGNILDRCPCSMENLVRIHCQPTC 122
QY 130 SPQSLFINVTRVAQGLQAPVAVAYEAFYQHSFAEQSYDSCSVRVPAATLAVGTC 189
DB 123 SPQAEFMHVATQNNKKGD-BYISSVDLHISTEYINKTKSCSVQSVPTQQLAFDLMC 181
QY 190 GYVSALCNAQRLNFGDGTGNGLAPLDTITFLLBPGQAVGSGIOPLNEGVARCNESQGD 249
DB 182 GAYSASRCNPTKWFNMGDATHPYVPFQIYIQHEP-KSNSNNFTPLNVTTPVCNQAVSS 240
QY 250 DVATSCQDCAASC-----PAIARPO-----ALDSTFYLGQMPGSLVLIILCSVPFV-VT 299
DB 241 KLPACSCSDCLSCPGPEPPRPFPFKIVGLDAYF-----VIMAAVFLVGL 288
QY 300 ILLVG---PRVAPARDKSNVDPKGTSLSDKLSFS-----THLLGFOFGQW 344
DB 289 VFIMGSLFTQSSMDNDNQVD---GNDVSDMPSENDSEYKEKGAHTETPLETFFTKW 345
QY 345 GTWVASWPLTILVSVIPVALAAGLVFTLTDVPELWSPNAPSQARSEKAFDQHGPF 404
DB 346 GYFASNPGLTLIAGASLVILGYGINFIEITTDVPLKWLASPNKSLRLEBEFDTKFSPF 405
QY 405 FRTNQVILTAPNRSSVRYDLSLLGNPNFSGILDLDLLELLEQLRHLQWSPQAQRN 464
DB 406 YRLEQIIKAVNLPQIVHNT-SNGPYTFGVPDREFLTAKVLDLQEGIKEINA-----NG 458
QY 465 ISLQDICVAPLNDNTSL--YDCCINSILQYQNNRTLLLTANQTLMGQTSQVDKQHP 522
DB 459 TQKIDICVAPLSDGSEIDVSCVQVSIWGYFGDRL-----DDHEDNGFNVTYLDAL 514
QY 523 LYCANAPLTFKDGTLALSCMADYGAPVPFPLAIGY-----KGKDYSEAEALIMTF 574
```

```

DB 515 YDCISNPLYL-----CLAPYGGPVDPAIALGGFLPPGQDLTGSTKTFELANAILTF 564
QY 575 SLNNYPAGDPRLAQAKLWEEAFLEEMRAF-QRRMAGMFQVTTFTABRSLEDEINRTABDL 633
DB 565 LVKNH-HNKTDLLENALTWEKKFVEFMTWYTKNNMSQYMDIAFTSERSIEDELNRSQSDV 623
QY 634 PIPATSYIVIFLYISLALGSYSSWSRVWDSKATILGCGVAVVLGAVMAAGFFSYLGR 693
DB 624 LTILVSYLIMPMYTAISLGHVKEFKRVFIDSKITLIGIGVITVLASVSSVGVFGYIGLP 683
QY 694 SSVLQVVPFELVLSVGADNIFIFVLEYQRIQPRPGEPREVHIGRALGRVAPSMLLCSLS 753
DB 684 ATLIIVEVIFVLAVGVDNIFILVQTHORQORPENETLEQQVGRILKVGPSMLLTSL 743
QY 754 EACFFFGALTPMPAVRTFALTSGLAVIDLFLQMSAFVALLSDSKQERASRLDVCCV 813
DB 744 ESFCFFLGLSDMPAVRAPALYAGVALIIDFLQLITCFVSLFTLDTKREERNMDICCFI 803
QY 814 KPQELPPPGQEGLLGFFQKAYAPFLHWTGRGVLLFLALFGVLSYSMCHISVGLDQ 873
DB 804 KGKPDSTSNEGLLYKFFSVYVFFLMKKIVRASVMVIFFAWLCSIAIAPRIDIGDQ 863
QY 874 ELALPKDSYLLDYFLFLNRYFEVGAPVYFVTTLTGYNFSEAGMNAICSAGCNFNFSFTQK 933
DB 864 ELAMPQDSFVLHYFQSLNENLNGPPVYFVLKGLDAYTNSSDQNLVCAGQYCNDDSVLTQ 923
QY 934 IOYATEFPPEQSYLAIPASSWDDFIDW-LTPSSCCRLYISGPNKDK--FCPSTVNSLNL 990
DB 924 IYLASRHSNQTYYIARPASSWIDDYFDMAAASCCCKY-----RKDSGDFCPH--QDTSCL 976
QY 991 KNCHSITMGS--VRPSVEQFHKVLPWFLNDRPNIKCPKGLAAYSTSVNLTSQGVL---A 1046
DB 977 R--CNITKNLSLRPEKEFVKILPFLFKDNPDCTCAKAGHAAYGAVRYSHSHERLINEA 1034
QY 1047 SRFMAYHKLKNSQDYTEALRAARELANITADLR-----KVP-GTDPAFEFVFFYITN 1099
DB 1035 SYFMAYHTLKSSADYFLALESARKISANITQMLQRLMSNGVPMASALTVEVFPYSVY 1094
QY 1100 VFYQYLTILPEGLFMLSLCLVPTFVAVSCLLGLDLRLSGLNLLSIVMLVDTVGFMALW 1159
DB 1095 VFYEQLTMTWSDTLQSMGISVLSIFVTVFLMGDFVHSALVVVITITMIVVNLGLMTYW 1154
QY 1160 DISYNAVSLINLAVSGMSVEFVSHITRSEFAISTKPTWLERAKEATISMGSAVPAGVAMT 1219
DB 1155 NISLNAVSLVNLWAVGISVEFCSHLVHSFATSKSVSQIDRAADSLSKMGSSIFSGITLT 1214
QY 1220 NLPGILVLGLAKAQLIQIFFFLNLILTLGLLHGLVFLPVILSYVGPDPVNPALAEQKR 1279
DB 1215 KPAFILVLAFAKSQIFQVYFPMYGLIIVIGAAHGLIFLPVLLSYIGAPVSNARLRYHSQ 1274
QY 1280 A-----EEAFAVM 1288
DB 1275 AAAREHETALAGIL 1287

RESULT 93
US-60-229-511-274
; Sequence 274, Application US/60229511
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; TITLE OF INVENTION: DRUG-METABOLIZING PHASE I PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000767
; CURRENT APPLICATION NUMBER: US/60/229,511
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 1525
; TYPE: PRF
; ORGANISM: HUMAN
```

US-60-229-511-274

Query Match 30.2%; Score 2087; DB 46; Length 1525;
Best Local Similarity 32.3%; Pred. No. 2.3e-173;
Matches 515; Conservative 237; Mismatches 461; Indels 382; Gaps 38;

QY 33 CAPYDECG-----KNPELSGLMTLSNVSCLSNTPARKITGDHLLILQKICPRLYTGP 85
DB 9 CWVYGECEIAYGDKRYNCEYSG-----PPKPLPKDG-----YDLSAGFFFG- 49

QY 86 NTOACCSAKQLVSLASLSITKALLTRCPACSDNFVNLHCHNTCSNQSLFNVTR---- 141
DB 50 NVSLCCDVRLQTLKDNQLPQLPQLSRCPSCFNLLNLFCELTCSRQSQFLNVATEDY 109

QY 142 VAQLGAGQAPVAVAYEAFYQHSFABQSYDCSRVRVPAATAVLAVGTMCVYGSAICNAOR 201
DB 110 VDPVNTQTKNVKELQYVVGQSFANAMYNACRDEAPSSNDKALGILGCKDADA-CNATN 168

QY 202 WLNFGQDTGNGLAPLDIT-----PHLEPGQAVSGIQLPNEGVARNCSQGGDVATC 254
DB 169 WTEYMFNKGNGQAPFTITPVFSDPPVH-----GMEPMNATKGCDSVDVETAPC 218

QY 255 SQDCAASCAPATARPO-----ALDSTFVLGQMPGSLVLIILICSPVAVVITLL---- 302
DB 219 SQDCSIVCGPKPQPPPPAPWTILGLDAMYIMMITYMAFLLVFFGAFVAVCWKRXRYF 278

QY 303 -----VGRFV-----APARD-----KSKXV-----DP 319
DB 279 VSEYTPIDSNIAFVNASDKVGLQWVRWCASAEATAKADWLVAFVAPSSQGGTAEP 338

QY 320 KIGTSLSDKLSFSTHTLLQPFQGGTGWVASHPLTILVLSVTPVVALAAGLVFTBLTTP 379
DB 339 RE-ASCCDPVSAFAFEGCLRLRTRWGSFCVRNPGCVIFPSLVFTACSGSLVFVRVTNP 397

QY 380 VELWAPNSOARSEKAFDOHGPRTNOVILTAPNRSSYDYDSLLGPK-NFSGILDJ 438
DB 398 VDLWAPSQARLEKEYPDQHGPFRTQELIIRAPLTDKHYQYPFGADVPPFGPLDI 457

QY 439 DILLLELLEQLRLHLQVMSPEAQNISLQDICYAPLNPNDTSLYDCCINSLLQYFQNNR 498
DB 458 QILHQVLDLQIAIEN--ITASVDNETVTLQDILCLAPLSYNT--NCTILSVLNFQNSH 512

QY 499 TLLLTANTQTLMGQTSQVDMKDHFLY----- 524
DB 513 SVLDHKKGGDDF---VIADYHTHFLYCVRVGDGEPWISAVDEGTQLLAAGHLTSLMPSP 569

QY 525 -----CAN----- 527

DB 570 IHSQGWICRSPSAPGVVADGLSENAGSLSTGQLSVCTALPVASCLGHCANISTYCSTC 629

QY 528 -----APLT 531

DB 630 FLFLCFRQTFHFAQAI FVKQLMTFLSRKSLIVPQCLQNKGNPNYLTGTESFYFAVCPLL 689

QY 532 FK-----DGTALAL-----SCMADYGAP 549
DB 690 FPDWPPPSGAPLRHAPCAIDGTGLYPASCVPILSSQAPASLNDTSLIHDPCLTGFGP 749

QY 550 VPPFLAIGYKGYSEBALMTWLSLNNYPAGDPLAQAKLWEAFLEEMKAFQRRMAG 609
DB 750 VPPWLVLGGYDQNNATLVITFPVNNYNDTEKLRQAQWEKEFINFVKNYKN---P 806

QY 610 MFQVFTFAERSLEDEINRTAEDLPIFATSYIVIFLXTLSLAGSYSSMSRVMVDSKATIG 669
DB 807 NLTISFTAERSLEDEINRESDDVTTVISYAIMFLYISLAUGHKSCRRLLVDKSVLSG 866

QY 670 LGGVAVLGVANMAAGFFSYLGIRSLVILQVVPFLVLSVGDADNIFI FVLEYQRLPRPG 729
DB 867 TAGILVLSSVACSLGVFSYGLPLTLVIEVIFPLVAVGVDNIFILVQAYQ----- 919

QY 730 EPREVHIGRALGRVAPSMMLCSLSBAICPFL-----GALTUPMVAVNTFALTGSL 778
DB 920 -----VYPHILTGLSICS-SFRDCEWLGISAFVSVTCGALSVMVAHVHTFSLFAGL 969

RESULT 94

US-60-212-657-129

; Sequence 129, Application US/6012657

; GENERAL INFORMATION:

; APPLICANT: Beasley, Ellen

; TITLE OF INVENTION: ISOLATED HUMAN PHASE I DRUG-METABOLIZING

; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN

; FILE OF INVENTION: DRUG-METABOLIZING PHASE I PROTEINS, AND USES THEREOF

; FILE REFERENCE: CLO00672

; CURRENT APPLICATION NUMBER: US/60/212,657

; CURRENT FILING DATE: 2000-06-19

; NUMBER OF SEQ ID NOS: 303

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 129

; LENGTH: 1529

; TYPE: PKT

; ORGANISM: HUMAN

US-60-212-657-129

Query Match 30.1%; Score 2083; DB 46; Length 1529;

Best Local Similarity 32.3%; Pred. No. 5.2e-173;

Matches 517; Conservative 236; Mismatches 457; Indels 392; Gaps 39;

Db 238 GYGINFIBITDVPKLVASPNKSRRLEREFDTKFSFFTRLEQIIIIKANVLPOIVNT-S 296
QY 427 LGPNFSGILDLLLELLELQRLHQLWSPQAQRNLSLODICYAPLNPNTSL--YD 484
Db 297 NGPYTGPVDFRFLTKVLDLQGIKEINA-----NGTQKXDCIYAPLSDGSEIDVQ 350
QY 485 CCINSLLOVFNQNRITLLLTANQTLMGQTSQVDWKDHFYCANAPUTFDGKTALASQMA 544
Db 351 CVVQSIWGWFGDDRRL-----DDHDEDNQFNVTYLDALYDCISNPYL-----CLA 396
QY 545 DYGAIPPFLAIGGY-----KGKDYSEARALIMTFSLNYPACDPRLAQKWEAF 596
Db 397 PYGPDVDPALAGGFLPPGDLTGSTKFPFLANAILITFLVKNH-HNKTDLENALTWEKFP 455
QY 597 LEMRAF-QRRMAGMQVFTFAERSLEDEINRTAEDLPIFATSYIVIFLYISLALGSYS 655
Db 456 VERMTYTKNNMSQYMDIAFTSERSIEDELNRESQSDVLTILVSYLIMFYAISLGHVK 515
QY 656 SWSRVWDSKATLGLGGVAVVLGAVNMGFFSYLGIRSLVLQVLPFLVLSVGDADNIF 715
Db 516 EFKRVFIDSKITLIGIGVVIIVLASVSVGVFGYIGLPATLIIVEIPFLVAVGVDNIF 575
QY 716 IFVLEYORLPREGEPREVHIGRALGRVAPSMLLCSLSAICFFLGALTMPAPVTRFALT 775
Db 576 ILVQTHORQKPNETLEQQVGRILKVGPSMLLTLSBSFCFELGGLSDMPAVRAFALY 635
QY 776 SGLAVILD FLOWMSAFVALLSLSKQZASRLDVCCVKPQELPPPGQGEGLLGFQKA 835
Db 636 AGVALIIDFLQITCFVSLFTLTKREERNMDICCFIKGKPDSTSNEGILYKFSV 695
QY 836 YAPFLHWTTRGVLLLFALFGVLSYMSCHISVGLDQBLALPKDSYLLDYFLNRYPE 895
Db 696 YVPFLMKKIVRASVMVIFFAWLCSIAIAPRIDIGLQBLAMPQDSFVLHYFQSLNENL 755
QY 896 VGAPVYFVTLGVNFSSEAGMNAICSSAGCNPFSTQKIQYATEPEEQSYLAIPASSWD 955
Db 756 IGPVYFVLKGLDAYTNSDQNLVCAQYCNDDSVLTQIYLASHRHSNQTYIARPASSWD 815
QY 956 DFLDM-LTPSSCRLYISGNKDK--FCPSTVNSLMLKNCMSITWGS-VRPSEVGFHY 1011
Db 816 DFDWAAASSCCKY-----RKDSGDFCPH-QDTSCLR-CNITKNSLLRBEKEFVKY 866
QY 1012 LPWFLNDRPNIKCPKGLLAAYSTSVNLTSBGQVL---ASRFMAHYHKLKNSQDYTBALRA 1068
Db 867 LPFPLKNDPDDTCAKAGHAAGVAVRNSHRLNTEASYFMAHYHTILKSSADYFLALES 926
QY 1069 ARELANIITADLR-----KVP-GTDPAPFVPPYTTNVPFYEQYLITPEGLFMLSCLV 1121
Db 927 ARKISANITQMLQGRLSNGVPMASALTVEVPFYSVYFVEYEQYLTWMSDTLQSMGISVL 986
QY 1122 PTFAVSCLLGLDLRGLLNLSIVMLVDVTGFMALWDISYNAVSLINLVSAGVMSVRF 1181
Db 987 SIPVTFVLMPDFVHSALVVITITWVNLGGLMTYWNISLNAVSLNVLNVAAGLSVRF 1046
QY 1182 VSHITRSPAISTKPTWLERAKEATISMGSAVAGVAMTNLPGILVLGLAKAQLIQIFFR 1241
Db 1047 CSHLVHSFATSKSVSQIDRAADSLRWGSSIFSGITLTKFAGILVLFAKSIQFQVYFR 1106
QY 1242 LMLLITLLGLLHGLVFLPVLISVGPDPVNPALAEOKRA-----EEAVAAM 1288
Db 1107 MYLGIYVIGAAHGLIFLPVLLSYIGAPVSNARLYHSQAAAEHETALAGIL 1157

RESULT 96

US-10-449-902-53619
; Sequence 53619, Application US/10449902
; GENERAL INFORMATION:
; APPLICANT: National Institute of Agricultural Sciences.
; APPLICANT: Bio-oriented Technology Research Advancement Institution.
; APPLICANT: The Institute of Physical and Chemical Research.
; APPLICANT: Foundation for Advancement of International Science.
; TITLE OF INVENTION: FULL-LENGTH PLANT cDNA AND USES THEREOF
; FILE REFERENCE: MOA-A0205YI-US

; CURRENT APPLICATION NUMBER: US/10/449,902
; CURRENT FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: JP 2002-203269
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: JP 2002-383870
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 56791
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 53619
; LENGTH: 1279
; TYPE: PRT
; ORGANISM: Oryza sativa
; US-10-449-902-53619

Query Match 27.5%; Score 1898.5; DB 34; Length 1279;
Best Local Similarity 34.0%; Pred. No. 7.7e-157;
Matches 450; Conservative 238; Mismatches 527; Indels 107; Gaps 31;

QY 3 EAGLGRWLLWALLRLLAQSEPYTTIHQPGYCAFYDECGKNPGLSGLMTLSNVCSLNTNP 62
Db 10 EANL---IFTAFVLLASLISQSAQAQ---CVMYGQCAKNDR-----GQPLNCGYNGP 56
QY 63 ARKITGDHLLILQKICPLRY-TGPNTQACCSAKQLVSLSEASLSITKALLTRCPACSDNFV 121
Db 57 PKPLDDEGLSILQKRCPHLVDSGDRATTCDDAAMLKBLDNVNAKVAGLFQRCPCIRNLL 116
QY 122 NLHCHNTCSNPQSLFINVTRVAQLGAGQLPAVVAVYEAIFYQHSFAEQSYDSCSRVRVPAAA 181
Db 117 GYWCMTCAPDOSLPMFVTKFGVGSSKV--VDANVNIHIDQEVLOATPFSCKGVWPFSSG 174
QY 182 TLAVGTCGVYGSALCNAQRWLNFGQDTGNGLAPLIDITFHLLPQGAQVSGIQPLNEGVA 241
Db 175 KZAMDMACGHAQAQSCTPQRFDFMGDPSPFPVFKMN---ITSSTGTDFDMDIPRRPKFV 231
QY 242 RCNQSQDDVATCSCODCAASC-ATARQALDSTFYLGOMPSLVLIIILCSFVAV 297
Db 232 SCSESFPNNSKPCVDCCKAPKTDGILLPPESKLKWSFLGL--NLIVIIDIGISAALI 289
QY 298 VTILLVGRVAPAR----DKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQGWGTWASWPL 353
Db 290 VAALLWALFAGQKWLDETDEITENK----LEKIGDSWSRFLERSFWLGTACAAHPV 345
QY 354 TILVLSVIPVVALAAGLVETLTDPPVELWSAPNSOARSEKAFHDHQHGFPTFNQVILT 413
Db 346 WVLFISSWIIAGLSYGVQPLVTTDPVQIWAAPNSLVROEKTFTDFPFPFVRAEQMFIK 405
QY 414 ANRNASRYDLSLLGPNFSGILDLLLELLELQRLHQLWSPQAQRNLSLODICYA 473
Db 406 AVGLDEVKHOT-PNGEITFGPVYNKEFLALARDLTMKITEL-----GKABGAGLETVCILA 459
QY 474 P-----LNPDNLSYDCCINSILQYFQNNRTLLLTANQTLMGQTSQVDWKDHFYCAN 528
Db 460 PLASDFLGPVTT--QCTVFNIWGFQNSLEFP-----EKGRNYLDQINKCLR- 506
QY 529 PLTFKDGDTALASCMDADYGAFFPFLAIGGY--KGK-----DYSEABALIMTFPSLNNY 579
Db 507 -----NAYDPDLCLASYGPAEPVIVGDYLEKQSVRDNADPKKARAAILTFVLRNS 559
QY 580 PAGDPRLAQAKLWEAFLEEMRAFQR-RMAGHFQVFTFAERSLEDEINRTAEDLPIFAT 638
Db 560 LNPD-EIEKAMKEWLYYDFMQNWTNTEKPDFMDVAFSERSVQDSIERVSKSEVSTVVI 618
QY 639 SYVIFLYISLALGSYSSWSRVWDSKATLGLGGVAVVLGAVNMGFFSYLGIRSLSLVI 698
Db 619 SYTVMFYLSIFSLSGRFTSWRTLVBSKMTLGGAGVVVWVSVVCAVGFGYIGTPTMLT 678
QY 699 LQVPFLVLSVGADNIFIFVLEYQRLPRPG-----EPREVHIGRALGRVAPSMLLCSLS 753
Db 679 VEVIPLVAVGVDNIFINVRTH-RMRDLSGLFSGVDTEARIGRTVGKAGPAVLSLSLS 737
QY 754 EACFPFLGALTMPAVRTPALTSGLAVILDFLLQMSAFVALLSLDSKQSEARLSDVCCCV 813
Db 738 ESACCLIGALSMPAVNTFALPAATALAINFLVQITFTFVALMTLDSNRVLRARLDVACCI 797

QY 1040 SD-----GOVLASFMAYHKPKNSQDYTRALRAARELANITADLRKVPGTDPAFVFPY 1095
DB 975 IDDVGMSTVQDSYFMOYSTTSTTSEBFYSQLREVRRISGEINAMFKE---NNVDABIFAY 1031
QY 1096 TITNVFEQVLTLPGLFMLSCLVPTTAVSCLLGLDLRSLGLNLLSVMLIVDTVGF 1155
DB 1032 CVFYIYEQVLTWGDAMFSLGSLVAIFLVLTLLITGLDITSTFVFLVVICILINMLGM 1091
QY 1156 MALWDSYNVAVSLINLVSAGMSVFEVSHITRSFALSTKPTMLERAKEATISMGSAVFPAG 1215
DB 1092 MWASINLNAISLVNLCVGVGFVAHVRSFK-RAEGTAQERARHSLNVTGSSVLG 1150
QY 1216 VAMTNLPGLVLGLAKAQIQLIPFFRLNLLITLLGLLHGLVFLVLTSLVGVDPNPAAL 1275
DB 1151 ITLTKFAGIWLGFNSQIFQVFFRMYLGIIVLIGAAHGLIILPVLLSLILGPPQK----L 1206
QY 1276 EQKRAEEAANVMVAS 1291
DB 1207 ARSSGAETASITITT 1222

RESULT 98

US-09-614-150A-2679
; Sequence 2679, Application US/09614150A
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/09/614,150A
; PRIOR FILING DATE: 2000-07-11
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-05
; PRIOR FILING DATE: 1999-10-19
; PRIOR FILING DATE: 1999-10-28
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-11-12
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 1999-12-28
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-01-12
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2679
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-09-614-150A-2679

Query Match 27.0%; Score 1864.5; DB 26; Length 1223;
Best Local Similarity 33.5%; Pred. No. 7.2e-154;
Matches 428; Conservative 237; Mismatches 498; Indels 113; Gaps 27;
QY 62 PARKITG-DHLLLOKICPRLY-----TGNP-TQACCSAKOLVSLSELSITKALLTRC 113
DB 14 PARPLNSPTSEIFAIFAKRCPLMYKEYKGESGDELSCCDAQIETWESGLSQADGVFSRC 73
QY 114 PACSDNFVNLHCHNTCSNPQSLFINVTRVLAQAGOLPAVVAAYEAFYQHSFAEQS----Y 169
DB 74 PTCRNMALTVCACTCAKNHTLFLTAINDTN-----DAGVDYVKYIDRLTDDTVSKIY 127
QY 170 DSCSRVRVPAANTLAVGTWCGVYGSALCNAQRWLNFGQDTGNGLAPLDTTFHLLBPGQAV 229
DB 128 NSCIGIQHTQTRGPAMDLCGCSYNKACTCNRYRWRWFPMGDSVGDYVFPQINRYKWSADAEE- 186

QY 230 GSGIQPLNEGVARCNCSEGDVVATCSQDCAASCPAIARPOALDSTFYLGQMPG--SLVL 287
DB 187 GSENIYLDLSPLKCBESYEDSYA-CACIDCEBSCPLTDAPTGDELWKIAGLVGVIFILA 245
QY 288 IILICSVFAVNTILLVGFVRVARDKSKWDPKGTSLSDKLSFSTHTLLGQF-----FQG 343
DB 246 LIACALSPFIIPWAGFKTSAP-----SVCNPTLFGFEPFYHGFR 285
QY 344 WGTWVASWPLTLLVLSVIVPVALAAGLVFELTDTDPVELWSAPNSQARSEKAFHDQHFGP 403
DB 286 WGTFCXKHPVIVLALCSWAIAGLSFGIRTWITTDPELVWAGEESTRIEKDITFDQHFGP 345
QY 404 FPRITQVILITAPNRRSYRYDLSLLGKPNFSGILDLLELLELELQERLRLHQLVMSPEAQR 463
DB 346 FYRTNQMFVAVNTQVFTHT-SNGVLNFGPAFYNFLAKVFELODSINKLGM-----AD 399
QY 464 NISLQDICYAP-LNPDNTSLYD-CCINSLLQVFNQNRITLLLTANTQTLGQTSQVQWKOH 521
DB 400 NEGLDKICYAPVLMAGETFTVDRCAIQSVYGYFQHDMDRF---ENSVDNSNNTYINLQ 456
QY 522 FLYCANAPLTFKGTALALSQADYGVFPFLAIGGY-----KGKDYSEAEALIMTFS 575
DB 457 LEDCLRVPM-MED-----CFGTFGGPIEPGLIAGVGMKPVAVGEDPDMLATGLVITFL 508
QY 576 LNNYPAGDPLRLAQAKLWEEAFLEEMRAFQORMMAGMFQVTTAERSLEDBINRTTADLPI 635
DB 509 GRNY-NDBSKLRPNMKWKLFDVFLRDYK---SDRLDIAYMAERSIQDAIVELSEGEVST 564
QY 636 PATSVIVPLIYLSLALGSYSWSRVWDSKATLGLGGVAVVILGVAANMAGPFSYLGRSS 695
DB 565 VVISYVWVYVAIALGHIRSCRGFLRESRIMLAIGGIVIVLASVVCISLGFMYGLDVTIT 624
QY 696 LVILQVPELVLSVGDNIFIVLEFORLPRPGEPRVHIGRALGRVAPSMLLCSLSSEA 755
DB 625 MLAEIVPEPLVLAGVDNIFIMVHTYQRLDHSKFTTHEAIGEAIGVQSPSLQTAGSEM 684
QY 756 ICFFLGALTMPDAVRTFALTSLGLAVILDFLQMSAFVALLSDSKQESRLDVCCKVK- 814
DB 685 ACFAIGCISDMPAVTKFMYAAIALLDFLQITAFVALMADEKRYLQDRLDMLCCVKS 744
QY 815 -----POELPPPGQEGLLGLGFQKAYAPFLHWTIRGVVILLFLALFGV 859
DB 745 GGGKINDEDEGCVDRPKEV-----GLLETFLFKNFPYSPFLSKPKVKSVLITFTVITCL 797
QY 860 SLYSMCHISVGLDQELALPKDSYLDYFLFLNRYFEVGVAPVYFVTTLGNFSPSEAGNMI 919
DB 798 SLWVTFSIEKGLDQEMSPKSHVVKYFRYMYVDLLAMGAPVYVWLKPGVLYSEPLQONLI 857
QY 920 CSSAGCNPFSFTQKIQYATEPPEQSYLAIPASSWVDDFDIMLTPSSCCRLYISGPNKDF 979
DB 858 CGGVECNNSLSVQLYTOAQYPEITSLARPASSWLDYIDMLAISDCCKYNT---TGGF 914
QY 980 CPSTVNSLNLKNCMSITMGSVRSPVEQPHKYLPMFLNDRPNIKCPKGLAAYSTSVNLT 1039
DB 915 CSSNSKSEDLPCERGFTEGLRPAETFNKYIPYFLPDLPAECAKAGASVADAVIT 974
QY 1040 SD-----GOVLASFMAYHKPKNSQDYTRALRAARELANITADLRKVPGTDPAFVFPY 1095
DB 975 IDDVGMSTVQDSYFMOYSTTSTTSEBFYSQLREVRRISGEINAMFKE---NNVDABIFAY 1031
QY 1096 TITNVFEQVLTLPGLFMLSCLVPTTAVSCLLGLDLRSLGLNLLSVMLIVDTVGF 1155
DB 1032 CVFYIYEQVLTWGDAMFSLGSLVAIFLVLTLLITGLDITSTFVFLVVICILINMLGM 1091
QY 1156 MALWDSYNVAVSLINLVSAGMSVFEVSHITRSFALSTKPTMLERAKEATISMGSAVFPAG 1215
DB 1092 MWASINLNAISLVNLCVGVGFVAHVRSFK-RAEGTAQERARHSLNVTGSSVLG 1150
QY 1216 VAMTNLPGLVLGLAKAQIQLIPFFRLNLLITLLGLLHGLVFLVLTSLVGVDPNPAAL 1275
DB 1151 ITLTKFAGIWLGFNSQIFQVFFRMYLGIIVLIGAAHGLIILPVLLSLILGPPQK----L 1206
QY 1276 EQKRAEEAANVMVAS 1291


```

; SEQ ID NO 514
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Drosophila
US-60-161-932-514

```

Query Match	27.0%	Score 1864.5;	DB 45;	Length 1223;
Best Local Similarity	33.5%;	Pred. No. 7.2e-154;		
Matches 428;	Conservative 237;	Mismatches 498;	Indels 113;	Gaps 27;

Qy	1040	SD----	GOVLASPMAYHKLKNSQDYTTLSALRAARELANITADLERKVFCTDPAFEVFPY	1099
Db	975	IDDVGMSTVQDSYFMOYSTTSTTSEEFYSQLREVRRISGEINAMFKE--NNVDASIFAY	1031	
Qy	1096	TITNPFVEOVLITLPGELFMWLSCLVPTFAVSCLLGLDLRSGLNLLSVMLITVTVGF	1155	
Db	1032	CVFPIYVEQYLTITWGDAMFSLGMSLVAIFLVTLLITGLDITSTFIVLFWVICLINLWGM	1091	
Qy	1156	MALWDLISYNAVSLINLNAVSGMSVPEVSHITTSFAISTKPTWLERAKEATISMGSAVAPG	1215	
Db	1092	MWMSLINLNAISLVNVLVVCVGLGVGFVAHIVRSFK-RAGETQAERARHSLNVTGSSVLSG	1150	
Qy	1216	VAMTNLPGLIVTGLAKAQIQLPFFPRNLNLLITLGLLHGLVFLPVLTLSVYGDVDPNPALAL	1275	
Db	1151	ITLTKFAGIVLWGFNSQIQFQVFYFMYLGIVLIGAHHGLIILLPVLLSLSLGPPQK----	1206	
Qy	1276	EQKRAEEAAVAAMVMS	1291	
Db	1207	ARSSGAEPITASITIT	1222	

Search completed: April 11, 2006, 00:56:31
Job time : 594 secs

```

; SEQ ID NO 514
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Drosophila
US-60-161-932-514

Query Match      27.0%; Score 1864.5; DB 45; Length 1223;
Best Local Similarity 33.5%; Pred. No. 7.2e-154;
Matches 428; Conservative 237; Mismatches 496; Indels 113; Gaps 27

QY      62  PARKITG-DHLILLOKICRPLY-----TGPN-TQACCSAKQLVSLNASLSITKALLTRC 113
DB      14  PARPLNSPTSEAFKRCRPMLYKEYKSGSGEDSLCCDAAQIETSSSGLSQADGVFSRC 73

QY      114 PACSDNFVNLHCHNTCSPNQSLFINVTRVAQLGAGQLPAVVAEAFYQHSFAEQS----Y 169
DB      74  PTCRNMALTVCAITCAKXHTLFLAYNDTN-----DAGVDYKYIDYRLTDDTVSKIY 127

QY      170 DCSRRVRVPAATLAVGTWCGYVGSGALCNAORWLNFGQDGTNGIAPLIDITPHLLEPQAV 229
DB      128 NSCIGIQHTQTRPAMDLCGSGYNAKTCNRYRWYEFMGVDGSDVVPQINYNKSEDAEE- 186

QY      230 GSGIQPLNEGVARCNESQDDVATCSQDCDCAASCAPAIARQALDSTFVLQMGQ--SLVL 287
DB      187 GSNEIYDLSPLKCGESYEDSYA-CACIDCBESCLTDAPTGPDELWKIAGLVGVTFILA 245

QY      288 IILCSVFPAVVTILLVGFVRPARDKSKMWDPKKGTSLSDKLSFTHTLLGQF----FQG 343
DB      246 LIIFACALSFFIIPWGAFGKTSAP-----SVCMTPLPGRFYFHGCFRI 285

QY      344 WGTWASWPLTLILSVIPVVALAAGLVFTBELTTPVELWSAPNSQARSEKAFHDQFGP 403
DB      286 WGTFCAKHEPVIVLALCSMAIAGLSFGIRYMTITTPVELWAGESQTRIEKQYDFDQHFGP 345

QY      404 EPRTNQVILTANRRSSRYVDSLLGPKNFGSILDDLILLELLELOERLRLHQLWSPEAQR 463
DB      346 FRTNQMFYKAVNQTYFTHET-SNGLVNFPGPAEFYFLKEVFLQDSIMKLGM-----AD 399

QY      464 NISLQDICYAP-LNPNTSLYD-CCINSILQYFQNNRTLLLTANQTLMGQTSQVQDKDH 521
DB      400 NEGLDKICYAPVLMAGEPTVDRCALQSVGYFQHDMDRP---ENSVDNSNTTYINVLNQ 456

QY      522 FLYCANAPLTFKDGITALALSCWADYGAPVPFFLAIGY-----KGKDYSEABALIMTFS 575
DB      457 LEDCLRVPM-MED-----CFGTGGPIEGIAVGGMPKVAVGEDPDYMLATGLVLITPL 508

QY      576 LNNYPAGDRLQAQKWEAFLEEMKRAFORRMAGMFGQVFTTARBSLEDEINRTTASDLPI 635
DB      509 GRNY-NDESKLEPNMKWEKLVDFLRDYY---SDRLDIAVMAERESIQDAIVELSEGEVST 564

QY      636 FATSIVIFLYISLALGSYSWSRVMVDSKATLGLGVAVVLGAVMAAMGFYSVLGRSS 695
DB      565 VVISYVMFVVAIALGHRISCRGFLRESRLMLAIGGIVIVLASVCSLGFNGYLDVTTT 624

QY      696 LVILQVPPVLVSGADNIFIPVLEYQRLPRRPGEPREHVHIGRALGRVAFPSMLCLSSEA 755
DB      625 MLAEIVPFLVAVGVDNIFIMWHTYQRLDHSKEKTTHEALGEAIGQVGPSILQTAGSEM 684

QY      756 ICFFLPGALTPMVAETFTALTSGLAVILDFLQMSAFVALLSLDSKQBSRLDVCVCKV- 814
DB      685 ACFAGICSDMPAVKTFPAMYAAIALLDFLLQITAFVALLMAIDEKRYLDGRDMLCCKVKS 744

QY      815 -----PQSLPPPGQEGLLGFFOKAVAPFLHLHWITRGVVLLFLALFGV 859
DB      745 GKKINDGDCGVDRPKEV-----GLLETFKKNFYSPFLLSKPKVKSULLITPTVITCL 797

QY      860 SLYSMCHISVGLDQELALPKDSYLLDYPLFLNRYFEYVGAPVYFVTTILGNPFSSEAGNNAI 919
DB      798 SLMWTPSYIEKGLDQSMSPKNSHVVKYPRYMWVDLLAMGAPVYVWLKPLNYSPLQONLI 857

QY      920 CSSACGNFSTFKIQYATEFPEQSYLAIPASSWVDDFIDWLTPSSCCRLYISGPNKDKP 979
DB      858 CGGVCECNNSLSVQLTYQAQYPEITSLRASSWLLDYIDWLAIISDCKKNVT---TGCF 914

```

THIS PAGE IS BLANK

GenCore version 5.1.7
 Copyright (c) 1993 - 2006 Bioacceleration Ltd.
 OM protein - protein search, using sw model
 Run on: April 11, 2006, 00:47:36 ; Search time 78 Seconds
 (without alignments)
 2379.065 Million cell updates/sec

Title: US-10-736-769-4
 Perfect score: 6909
 Sequence: 1 MAEAGLRGWLWALLRLAQ.....GSIKAGATSNFLPNNGRQF 1332

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 592492 seqs, 139314628 residues

Total number of hits satisfying chosen parameters: 592492

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Pending Patents AA New:*

- 1: /SIDSS/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /SIDSS/ptodata/2/paa/US06_NEW_COMB.pep.*
- 3: /SIDSS/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /SIDSS/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /SIDSS/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /SIDSS/ptodata/2/paa/US10_NEW_COMB.pep.*
- 7: /SIDSS/ptodata/2/paa/US11_NEW_COMB.pep.*
- 8: /SIDSS/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	6909	100.0	1332	6	US-11-301-094-6
2	6872.5	99.5	1359	6	US-11-270-796-22
3	5407	78.3	1333	6	US-11-270-796-3
4	2402.5	34.8	1278	1	PCT-US06-05584-822
5	2402.5	34.8	1278	6	US-11-191-274A-129
6	2402.5	34.8	1278	6	US-11-191-274A-130
7	2402.5	34.8	1278	7	US-11-385-692-2400
8	2402.5	34.8	1278	7	US-11-385-692-2401
9	1341.5	19.4	1383	6	US-11-301-094-2
10	1046	15.1	1274	6	US-11-301-094-4
11	782.5	11.3	419	6	US-10-953-349-21065
12	588.5	8.5	967	6	US-10-461-673-16739
13	583.5	8.4	1182	6	US-11-332-764-2
14	583.5	8.4	1182	6	US-11-337-244-149
15	482	7.0	891	8	US-60-772-265-1197
16	455	6.6	1358	8	US-60-772-265-241
17	437	6.3	831	7	US-11-360-355-133805
18	349	5.1	783	6	US-10-461-673-16750
19	325	4.7	204	7	US-11-360-355-141351
20	261	3.8	575	7	US-11-360-355-120785
21	252	3.6	505	6	US-10-461-673-16871
22	229.5	3.3	465	7	US-11-360-355-141367
23	225	3.3	542	6	US-11-214-063A-1670
24	222.5	3.2	422	7	US-11-360-355-141365
25	219	3.2	1137	6	US-10-461-673-16889

Sequence 10541, A	1330	3.2	219	6	US-10-461-673-10541
Sequence 141352, A	126	2.8	190	6	US-11-360-355-141352
Sequence 3548, Ap	632	2.2	152.5	7	US-11-293-697-3548
Sequence 127069, A	164	2.0	141	7	US-11-360-355-127069
Sequence 120169, A	484	2.0	139.5	7	US-11-360-355-120169
Sequence 140206, A	252	2.0	135.5	7	US-11-360-355-140206
Sequence 7050, Ap	808	1.8	126	6	US-10-724-972B-7050
Sequence 149589, A	171	1.8	123.5	7	US-11-360-355-149589
Sequence 20, Appl	1043	1.8	121.5	6	US-10-536-606-20
Sequence 2044, Ap	768	1.7	118.5	6	US-11-203-806A-11
Sequence 1692, Ap	788	1.7	118.5	6	US-11-214-063A-2044
Sequence 1828, Ap	619	1.7	117.5	8	US-60-732-162-1828
Sequence 46, Appl	619	1.7	116.5	6	US-11-312-958-46
Sequence 12167, A	619	1.7	116.5	6	US-10-461-673-12167
Sequence 45206, A	488	1.7	114	8	US-60-752-355-45206
Sequence 2874, Ap	985	1.7	114	6	US-11-293-697-2874
Sequence 226, App	697	1.6	112.5	6	US-10-703-799B-226
Sequence 152528, A	201	1.6	111	7	US-11-360-355-152528
Sequence 21, Appl	958	1.6	111	8	US-60-781-953-21
Sequence 2027, Ap	494	1.6	110.5	6	US-11-045-004-2027
Sequence 2398, Ap	451	1.6	109.5	6	US-11-045-004-2398
Sequence 73134, A	758	1.6	109.5	7	US-11-056-355B-73134
Sequence 73132, A	827	1.6	109.5	7	US-11-056-355B-73132
Sequence 33716, A	3979	1.6	109.5	8	US-60-752-355-33716
Sequence 39891, A	474	1.6	108	8	US-60-752-355-39891
Sequence 5410, Ap	901	1.5	107	6	US-10-724-972B-5410
Sequence 1237, Ap	629	1.5	106.5	6	US-10-567-867-1237
Sequence 16314, A	435	1.5	106	8	US-60-752-355-16314
Sequence 3216, Ap	924	1.5	106	7	US-11-027-399A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-028-099A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-027-878A-3216
Sequence 3216, Ap	924	1.5	106	7	US-11-028-291A-3216
Sequence 15654, A	637	1.5	105.5	6	US-10-917-503B-15654
Sequence 272, App	783	1.5	105.5	8	US-10-703-799B-272
Sequence 19639, A	308	1.5	105	7	US-11-360-355-19639
Sequence 69789, A	1080	1.5	105	7	US-11-056-355B-69789
Sequence 69788, A	1097	1.5	105	7	US-11-056-355B-69788
Sequence 129465, A	122	1.5	104	7	US-11-056-355B-129465
Sequence 149551, A	238	1.5	104	6	US-11-360-355-149551
Sequence 4902, Ap	510	1.5	104	6	US-10-724-972B-4902
Sequence 191, App	692	1.5	104	6	US-10-159-257C-191
Sequence 69565, A	757	1.5	104	7	US-11-056-355B-69565
Sequence 69564, A	771	1.5	104	7	US-11-056-355B-69564
Sequence 8918, Ap	826	1.5	103.5	8	US-11-056-355B-8918
Sequence 7153, Ap	618	1.5	103.5	1	PCT-US06-00964-7153
Sequence 25099, A	618	1.5	103.5	6	US-11-330-403-7153
Sequence 4330, Ap	759	1.5	103.5	8	US-60-752-355-25099
Sequence 246, App	1136	1.5	103	6	US-11-045-004-1583
Sequence 37889, A	463	1.5	103	8	US-60-752-355-4330
Sequence 1806, Ap	470	1.5	102.5	6	US-10-703-799B-246
Sequence 208, App	472	1.5	102.5	6	US-60-752-355-37889
Sequence 25298, A	538	1.5	102.5	8	US-10-498-451-1806
Sequence 14490, A	752	1.5	102.5	8	US-60-742-219-208
Sequence 14490, A	304	1.5	102	1	PCT-US06-07642-14490
Sequence 22441, A	304	1.5	102	8	US-60-658-984A-14490
Sequence 1750, Appl	569	1.5	102	8	US-60-752-355-22441
Sequence 41, Appl	1364	1.5	101.5	6	US-10-498-451-1750
Sequence 12928, A	634	1.5	101.5	6	US-60-781-306-41
Sequence 2249, Ap	397	1.5	101	6	US-10-917-503B-12928
Sequence 260, App	397	1.5	101	6	US-10-533-519-2249
Sequence 260, App	475	1.5	101	6	US-10-206-921A-260
Sequence 260, App	475	1.5	101	6	US-10-184-614A-260
Sequence 1676, Ap	475	1.5	101	8	US-10-184-615A-260
Sequence 7829, Ap	556	1.5	101	8	US-60-752-355-7829
Sequence 3734, Ap	607	1.5	101	7	US-11-174-307B-3734
Sequence 19, Appl	1016	1.5	101	1	PCT-US05-46487-19

99 101 1.5 1016 6 US-11-314-892-19 Sequence 19, Appl
100 101 1.5 1016 7 US-11-375-551-16 Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-11-301-094-6
; Sequence 6, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 6
; LENGTH: 1332
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-301-094-6

Query Match 100.0%; Score 6909; DB 6; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1332; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAEAGLGGWLLALLRLAQSBPYTHIQPGYCAFYDECGKPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLGGWLLALLRLAQSBPYTHIQPGYCAFYDECGKPELSGLMTLSNVSCLSN 60
QY 61 TPARKITGDLHLLQKICRLPYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDLHLLQKICRLPYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120
QY 121 VNLHCHTCSNQSGLFINTRVAQLGAGQLPAVVAEYFQHSFABQSDYSDCSRVRPAA 180
DB 121 VNLHCHTCSNQSGLFINTRVAQLGAGQLPAVVAEYFQHSFABQSDYSDCSRVRPAA 180
QY 181 ATLAVGTMCGVGSALCNQRLNFGDGTGNGLAPLDITPHLLEPGQAVGSGIQPLNEGV 240
DB 181 ATLAVGTMCGVGSALCNQRLNFGDGTGNGLAPLDITPHLLEPGQAVGSGIQPLNEGV 240
QY 241 ARCNESSQDDVATCSQDCAASCAPAIARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESSQDDVATCSQDCAASCAPAIARQALDSTFYLQMPGSLVLIILCSVFAVVTI 300
QY 301 LLVGFVAPARDKSKWDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGFVAPARDKSKWDPKKGTSLSDKLSFSFTHLLGQFFQCGWGTWVASWPLTILVLSV 360
QY 361 IPVVALAAGLVFTELTDPVELMSAPNSQARSEKAFHDFGPFPTNVLITAPNRSY 420
DB 361 IPVVALAAGLVFTELTDPVELMSAPNSQARSEKAFHDFGPFPTNVLITAPNRSY 420
QY 421 RYDSSLGPKNFGSGLDLDLLELLEQLERLHLQVWSPQARNISLQDICVAPLNPDT 480
DB 421 RYDSSLGPKNFGSGLDLDLLELLEQLERLHLQVWSPQARNISLQDICVAPLNPDT 480
QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTKDGTALAL 540
QY 541 SCHADYGAPVFPPLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCHADYGAPVFPPLATGGYKGDYSEAEALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEM 600
QY 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660

DB 601 RAFORMMAGMFQVTTAERSLEDEINRTTAEDLPFATSYIVIFLYISLALGSYSWSRV 660
QY 661 MYDSKATILGGVAVVILGAVMAAGFFSYILGRSSLVILQVVPFLVLSVGADNIFIVLE 720
DB 661 MYDSKATILGGVAVVILGAVMAAGFFSYILGRSSLVILQVVPFLVLSVGADNIFIVLE 720
QY 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
DB 721 YQRLPRRPGEPREVHIGRALGRVAPSMLLCSLSEAI CFFLGALTTPMPAVRTFALTSLGLAV 780
QY 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCCKPQDELPPPGQGGGLLGLFGKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQEAASRLDVCCCKPQDELPPPGQGGGLLGLFGKAYAPFL 840
QY 841 LHMITRGVVLLFLALFGVSLYSMCHI SVGLDOELALPKDSYLLDYFLFLNRYFEVGPV 900
DB 841 LHMITRGVVLLFLALFGVSLYSMCHI SVGLDOELALPKDSYLLDYFLFLNRYFEVGPV 900
QY 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSTFKIQYATEFFPQSYLAIPASSWVDDFIDW 960
DB 901 YFVTTLGYNFSSEAGMNAICSSAGCNFSTFKIQYATEFFPQSYLAIPASSWVDDFIDW 960
QY 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNMSITMGSVRPSVEQFHKYLPMFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDFCPSVTNSLNCNMSITMGSVRPSVEQFHKYLPMFLNDRP 1020
QY 1021 NIKCPKGLAAYSTSNLTSDQVLSRPMAYHKPKNSODYTEALRAARELAANITADL 1080
DB 1021 NIKCPKGLAAYSTSNLTSDQVLSRPMAYHKPKNSODYTEALRAARELAANITADL 1080
QY 1081 RKVPGTDPAFEVFPYTTITNVFYBOYLTITLPEGLPMLSCLVPTFAVSCILLGLDLSGGL 1140
DB 1081 RKVPGTDPAFEVFPYTTITNVFYBOYLTITLPEGLPMLSCLVPTFAVSCILLGLDLSGGL 1140
QY 1141 NLLSIVMLVDTVGPMAWDISYNVAVSLINLVSAGMSVEFVSHITRSPATSKPTWLER 1200
DB 1141 NLLSIVMLVDTVGPMAWDISYNVAVSLINLVSAGMSVEFVSHITRSPATSKPTWLER 1200
QY 1201 AKEATISMGSAVAGVAMTNLPGIILVGLAKAQLIQIFFRNLALTLLGLLHGLVFLPV 1260
DB 1201 AKEATISMGSAVAGVAMTNLPGIILVGLAKAQLIQIFFRNLALTLLGLLHGLVFLPV 1260
QY 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
DB 1261 ILSYVGPDPVPALALEQKRAEAAVAVMVASCPNHPSPRVSTADNIYVNHSPFGSIKGAGA 1320
QY 1321 ISNFLPNNGRQF 1332
DB 1321 ISNFLPNNGRQF 1332

RESULT 2

US-11-270-796-22
; Sequence 22, Application US/11207096
; GENERAL INFORMATION:
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22
; LENGTH: 1359
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-270-796-22
Query Match 99.5%; Score 6872.5; DB 6; Length 1359;
Best Local Similarity 97.9%; Pred. No. 0;

Matches 1330; Conservative 0; Mismatches 2; Indels 27; Gaps 1;

QY 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60
DB 1 MAEAGLRGWLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSN 60

QY 61 TPARKITGDLHLLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120
DB 61 TPARKITGDLHLLILLOKICPRLYTGPNTOACCSAKQLVSLASISITKALLTRCPACSDNF 120

QY 121 VNLHCHNTCPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAA 180
DB 121 VNLHCHNTCPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAA 180

QY 181 ATLAVTGTCGVGYSALCNAQRMWLNFGDGTGNGLAPLDTIFHLLEPQAVGSGIQLNEGV 240
DB 181 ATLAVTGTCGVGYSALCNAQRMWLNFGDGTGNGLAPLDTIFHLLEPQAVGSGIQLNEGV 240

QY 241 ARCNESQDDVATCSQDCODCAASCPAIPALDSTFVLGQPGSLVLIILCSVFAVVTI 300
DB 241 ARCNESQDDVATCSQDCODCAASCPAIPALDSTFVLGQPGSLVLIILCSVFAVVTI 300

QY 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360
DB 301 LLVGRVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 360

QY 361 IPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSY 420
DB 361 IPVVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSY 420

QY 421 RYDSSLILGPNFSGIILDLLELLELRLHQLVMSPEAQNSIQDICYAPLNPDNT 480
DB 421 RYDSSLILGPNFSGIILDLLELLELRLHQLVMSPEAQNSIQDICYAPLNPDNT 480

QY 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKQHFYCANAPLTKDGTALAL 540
DB 481 SLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDKQHFYCANAPLTKDGTALAL 540

QY 541 SCWADYGAPVFPFLAIGGYKDYSEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEM 600
DB 541 SCWADYGAPVFPFLAIGGYKDYSEALIMTFLSNYPAGDPRLAQAKLWEEAFLEEM 600

QY 601 RAFORMAGMFWQVTFPABRSLEDEINRTTAEDLPFATSVIIVIFLYISALGSYSWSRV 660
DB 601 RAFORMAGMFWQVTFPABRSLEDEINRTTAEDLPFATSVIIVIFLYISALGSYSWSRV 660

QY 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLVILQVPPFLVLSVAGADNIFIFVLE 720
DB 661 MYDSKATLGLGVAVVLGAVMAAMGFFSYLGRSSLVILQVPPFLVLSVAGADNIFIFVLE 720

QY 721 YORLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780
DB 721 YORLPRRPGEPREHVHIGRALGRVAPSMMLCSLSEALCFPLGALTMPAVRTFALTSLGLAV 780

QY 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGELLGFFQKAYAPFL 840
DB 781 ILDFLLQMSAFVALLSLDSKQREASRLDVCCVKPQELPPPGQEGELLGFFQKAYAPFL 840

QY 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900
DB 841 LHWITRGVLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYFEVGAPV 900

QY 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSTQKIQYATFPPQSYLEIAPASSWVDDFIDW 960
DB 901 YFVTTILGYNFSSEAGNNAICSSAGCNPFSTQKIQYATFPPQSYLEIAPASSWVDDFIDW 960

QY 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNLCNMSITMGSRVPSVBOFHKLIPWFLNDRP 1020
DB 961 LTPSSCCRLYISGPNKDKFCPSVNSLNCNLCNMSITMGSRVPSVBOFHKLIPWFLNDRP 1020

QY 1021 NIKCPKGGGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053
DB 1021 NIKCPKGGGLAAYSTSVNLTSDGQVL-----ASRFMAYH 1053

QY 1054 KPLKNSQDYTBALRAARELAANITADLRKVPGTDPAPFVFPYTTITNVFEQYLTILPEGL 1113
DB 1081 KPLKNSQDYTBALRAARELAANITADLRKVPGTDPAPFVFPYTTITNVFEQYLTILPEGL 1140

QY 1114 FMLSCLVPTFAVSCLLIGLRLSGLNLLSVMLIVDTVGMALWDISYNAVSLNLVLS 1173
DB 1141 FMLSCLVPTFAVSCLLIGLRLSGLNLLSVMLIVDTVGMALWDISYNAVSLNLVLS 1200

QY 1174 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1233
DB 1201 AVGMSVEFVSHITRSPFAISTKPTWLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQ 1260

QY 1234 LIQIFFFFRLNLITLLGLLHGLVFLPVILSYVGPDPVPALAEQKAEBAVAAMVASCP 1293
DB 1261 LIQIFFFFRLNLITLLGLLHGLVFLPVILSYVGPDPVPALAEQKAEBAVAAMVASCP 1320

QY 1294 NHPSRVSTADNLYVNHSPSGSIKAGATSNFLPNNGRQF 1332
DB 1321 NHPSRVSTADNLYVNHSPSGSIKAGATSNFLPNNGRQF 1359

RESULT 3
US-11-270-796-3
; Sequence 3, Application US/11270796
; GENERAL INFORMATION:
; APPLICANT: Dong, Jianli
; TITLE OF INVENTION: TREATMENT OF CANCER BY SIMULTANEOUS INHIBITION OF BRAF AND
; TITLE OF INVENTION: RESTORATION OR MIMICRY PP p16 Ink4a ACTIVITY
; FILE REFERENCE: 02420/1201581-US1
; CURRENT APPLICATION NUMBER: US/11/270,796
; CURRENT FILING DATE: 2005-11-08
; PRIOR APPLICATION NUMBER: 60/592,592
; PRIOR FILING DATE: 2004-07-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 3
; LENGTH: 1333
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-270-796-3

Query Match 78.3%; Score 5407; DB 6; Length 1333;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 1028; Conservative 126; Mismatches 174; Indels 4; Gaps 4;

QY 4 AGLRGWLLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA 63
DB 3 AAWQGWLLWALLRLAQSEPYTHIQGYCAFYDECGKNPELSGLMTLSNVSCLSNTPA 62

QY 64 RKITGDHLLILLOKICPRLYTGPN-TOACCSAKQLVSLASISITKALLTRCPACSDNFVN 122
DB 63 RHVTGDHLLILLOKICPRLYTGPN-TOACCSAKQLVSLASISITKALLTRCPACSDNFVN 122

QY 123 LHCHNTCPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAAAT 182
DB 123 LHCHNTCPNOSLFINTVRAQLGAGOLPAVVAEAFYQHSFABQSVDSCSRVRPAAAT 182

QY 183 LAVGTGCGVYGSALCNAQRMWLNFGDGTGNGLAPLDTIFHLLEPQAVGSGIQLNEGVAR 242
DB 183 LAVGTGCGVYGSALCNAQRMWLNFGDGTGNGLAPLDTIFHLLEPQAVGSGIQLNEGVAR 242

QY 243 CHNESQDDVATCSQDCODCAASCPAIPALDSTFVLGQPGSLVLIILCSVFAVVTILL 302
DB 243 CHNESQDDVATCSQDCODCAASCPAIPALDSTFVLGQPGSLVLIILCSVFAVVTILL 302

QY 303 VGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 362
DB 303 VGRFVAPARDKSKMVDPKKGTSLSDKLSFSTHTLLGQFFQCGWGTWVASWPLTILVLSV 362

QY 363 VVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSYRY 422
DB 363 VVALAAGLVFTELTTPVELWSAPNSQARSEKAFHQHFGPFFRTNOVILTAPNRSYRY 422

```
QY 423 D5LLGPKNFSGILDILLLELLELOERLRHLQVMSPEAQRNLSLODICYAPLNDNTSL 482
Db 423 D5LLGPKNFSGILSDLLQELLELOERLRHLQVMSHEAQRNLSLODICYAPLNPHNTSL 482
QY 483 YDCCNSLSLQYFQNNRTLLLTANQTLMGQTSQVDWKDHFLYCANAPLTFKDGTLALASC 542
Db 483 TDCCVNSLSLQYFQNNHTLLLTANQTLNGQTSVDWKDHFLYCANAPLTYKDGTLALASC 542
QY 543 MADYGAPVPPFLAIIGYKGKDYSEABALIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRA 602
Db 543 IADYGAPVPPFLAVGQYQGTYSABALITFSINNYPADDPMAHAKLWBEAFLEEMQS 602
QY 603 FORRMAGHPOVTFEATERSLEDEINRTAEDLPIFATSYIVIELYISLALGSYSSSRVAV 662
Db 603 FORSTADKFOIAFSAERSLEDEINRTIQDLVFAISLIVFLYISLALGSYSSSRVAV 662
QY 663 DSKATILGCGVAVLGVMAAMGFFSYLIGIRSSVLQVPPFLVLSVGNADNIFIFVLEVQ 722
Db 663 DSKATILGCGVAVLGVMAAMGFFSYLIGVPSLVIQVPPFLVAVGADNIFIFVLEVQ 722
QY 723 RLPRRPGEPREVIHGRALGRVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSGLAIVL 782
Db 723 RLPRMPGEOREAHIGRTLGSVAPSMMLCSLSBAICFFLGALTPMPAVRTFALTSGLAIF 782
QY 783 DELLQMSAFVALLSDSKRQEARLDVCCVXPQELPPGQCEGILLGFFQKAYAPFLH 842
Db 783 DELLQMTAFVALLSDSKRQEARSDVCCFSRNLPPPKQEGELLGFFRKIYTPFLH 842
QY 843 WTRGWVLLFLFALGVSILYSCHISVGLDQDLALPKDSYLLDYFLFLNRYEVEGAPVYF 902
Db 843 RIRPVLVLLFLVFGANILMCNISVGLDQDLALPKDSYLLDYFLFLNRYLEVGPPVYF 902
QY 903 VTLTGVNFSSEAGMAICSSAGCNFSPQKIQYATEPPEQSYLAIPASSWVDDPIDMT 962
Db 903 DTTSGYNFSTEAGMAICSSAGCESFSLTKIQAISEFPNQYVAIAASSWVDDPIDMT 962
QY 963 P-SSCCRLYISGPNKDCPSTVNSLNCNKMSITMGSVRPSVQFHKYLPWFINDRPN 1021
Db 963 PSSCCRIYTRGPHKDEPCPSTDTFNCUKNCNRTLGPVRTTQFHKYLPWFINDTPN 1022
QY 1022 IKCPKGLAAYSTVNLTSQGVLASRFMAVYHKPLKNSQDYTEALRAARELAANITADIR 1081
Db 1023 IRCPKGLAAYSTVNLSDGQIIASQFWAYHKPLRNSQDTEALRASLLAANITAEIR 1082
QY 1082 KVPGTDPAPREVPPYITNVFYQYLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLLN 1141
Db 1083 KVPGTDPNPEVPYITNSVYQYLTVLPGLFTLALCFVPTFVVCYLLGLDIRSGILN 1142
QY 1142 LLSIVMILVDTGFWALWDISYNAVSLINLAVAGMSVEFVSHITRSPAISTKPTWLER 1201
Db 1143 LLSIIMILVDTIGLMAVWIGSYNAVSLINLAVAGMSVEFVSHITRSPAISTKPTLER 1202
QY 1202 KEATISMSGSAVAGVAMTNLPGILVLGLAKAQLIQIFFPRMLLITLLGLHGLVLPVI 1261
Db 1203 KDATIFMGSAVAGVAMTNFPGILVLGLAQALIQIFFPRMLLITLLGLHGLVLPV 1262
QY 1262 LSYVGDVNPALALEQKRAEEAAVAVMVASCPNHPRSRVSTANIIYNHSGFEGS-1KG 1320
Db 1263 LSYLGDVNPQALVLEKLEATEA-AMVSEPCQYPPPADANTSDYVNFNPEFPEINA 1321
QY 1321 ISNFLPNNGROF 1332
Db 1322 ASSSLPKSDQKF 1333
```

RESULT 4

```
PCT-US06-05584-822
; Sequence 822, Application PC/TUS0605584
; GENERAL INFORMATION:
; APPLICANT: DANA-FARBER CANCER INSTITUTE, INC.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION,
; TITLE OF INVENTION: ASSESSMENT, PREVENTION, AND THERAPY OF CANCER
```

```
; FILE REFERENCE: DFS-064.25(25992-6425)
; CURRENT APPLICATION NUMBER: PCT/US06/05584
; CURRENT FILING DATE: 2006-03-02
; PRIOR APPLICATION NUMBER: 60/690,064
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: 60/654,227
; PRIOR FILING DATE: 2005-02-17
; NUMBER OF SEQ ID NOS: 848
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 822
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US06-05584-822
```

```
Query Match 34.8%; Score 2402.5; DB 1; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGMILLALLRLAQSEPYTTIHQPCYCAFYDECG-----KNPELSGLMTLSNVSCLS 59
Db 4 RGLALGILLLLLCAPQFSQ-----SCVWYGECCGAYGDKRYNCBYSG-----46
QY 60 NTPARKITGDLHLLILKICPRLYTGPNTOACCSAKQLVSLSEASLSITKALLTRCPACSDN 119
Db 47 --PKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRLQTLKDNLQLPQLQFLSRCSFCFVN 103
QY 120 FVNHLCHNTSPNOSLFINVTR----VAQLGAGQLPAVVAEAFYQHSFASQSYDSCSRV 175
Db 104 LNLFCELTCSPROSFLNVTATEDYDVPVNTQNTKNVKEQLQYVVGOSFANAMYNACRDV 163
QY 176 RVPAANTLAVTGMCGVYGSALCNAORWLNFGDGTNGLAPLDIT-----PHILLEPGA 228
Db 164 EAPSNDKALGILCKODADA-CNATNWEYMFNKONGQAPFTITPVFSDFPVH-----215
QY 229 VSGSIQPLNEGVARCNESQDDVATCSQDCAASCFAIARPO-----ALDSTFYLG 279
Db 216 ---GWEPNNATKGCDESVDVETAPCSQDCSIVCGPKQPPPPAPWTLGLDAMVIM 272
QY 280 QMPGSLVLIILCSVFAVVTILL-----VGFVPAAPADKSKWDPKGTSL 326
Db 273 WITYMAFLVFPFAFFVWVCYRKRYFVSEYTPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLGGPQFGWTVASWPLTILVSLVPPVVALAAGLVFTELTTPDVELWSAP 386
Db 325 DPVSAAPFGCLRLRLETRWGSFCVRNPGCVIFESLVFITACSSGLVFRVVTNPDLWSAP 384
QY 387 NSQARSEKAFDQHFPPFRTNQVILTAPNRSYRYSLLGLPK-NFGSILDLDLLELL 445
Db 385 SSQARLEKEYFDQHFPPFRTTEQLIIRAPLTDKHIYQVPVSGADVFPGLDIQLHQVL 444
QY 446 ELQERLRHLQVMSPEAQRNLSLODICYAPLNDNTSLDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN-ITASYDNETVTLQDCLAPLSPYNT---NCTLSVLNPFQNSHSLDHKK 499
QY 506 NOTLMGQTSQVDWKDHFLYCANAPLTFKDGTLALASCMAVYCAPVFPFLAIGYKGKDY 565
Db 500 GDDFF---VYADITHFLYCVRAPASLNDTSLHDFCLGTGFGPFPVPLVGLGYDDQYN 556
QY 566 EBAELIMTFSLNYPAGDPRLAQAKLWBEAFLEEMRAFQRRMAGMFQVTFEATERSLEDE 625
Db 557 NATALVITFPVNNYNDTEKLQRAQWKEKFINFKVKNYKN---PNLTISFTEATERSLEDE 613
QY 626 NRTTAEDLPIFATSYIVIELYISLALGSYSSSRVMSKATLGLGVAVLGVMAAMG 685
Db 614 NRESDSVFTVVISYAIMFLYISLALGHKSCRRLLVDSKSLGSLAGILVLSVACSJG 673
QY 686 PFSYIGIRSSVLQVPPFLVLSVGNADNIFIVLEVYQRLPRRPGEPREVIHGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFPLVAVGVDNIFILVQAYQORDERLQGETLDQQLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFLGALTPMPAVRTFALTSGLAIVLQFLLQMSAFVALLSDSKRQEAR 805
```

```
Db 734 SMFLSPSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLLOITCFVSLLLGLDIIKQEK 793
Qy 806 RLDVCCVKPQELPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLALFGVSLYS 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLKDKWMRPVIAIFVGLSPFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLPLNRIFEVGAPVYFVTLGYNSSSEAGNAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMDYFYSQYLHAGPPVYFVLEBEGHDYTSKQGNMVGCGMG 913
Qy 925 CNMFSTQKIQVATEPPEQSYLAIPASSWDDFIDWLT-SSCCRLYISGPNKDFCPS 983
Db 914 CNNDSLVQOIFNAQOLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSVRPSVEQFHKYLPWFILNDRPNIKCPKGLGAAYSYSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPBEGKQRPQGGDFMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DGOVLASRFWAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFVPPYITNV 1100
Db 1030 GTRVGATYPMYTHVLTQTSADFDALKARLIASNT-ETMGINGS--AIRVFPYSVFTV 1086
Qy 1101 FYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIYVILVDTVGFALWD 1160
Db 1087 FYEQLTIIIDTIFNLGVSIGAIPLVTVWLLGCELWSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLERAKEATISMGSAVPAVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKGSRAERAEALAHMSSVFSGITLTK 1206
Qy 1221 LPEGLVGLAKAQLOIIFPRLMLLTILGLLHGLVFLPVLVSYGPDVNP 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLAMVLLGATHGLIFLPVLLSYIGPSVKA 1258

RESULT 5
US-11-191-274A-129
; Sequence 129, Application US/11191274A
; GENERAL INFORMATION:
; APPLICANT: Applera Corporation
; APPLICANT: Bruno DOMO
; TITLE OF INVENTION: LUNG DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CLO01536PROV
; CURRENT APPLICATION NUMBER: US/11/191,274A
; CURRENT FILING DATE: 2005-07-28
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 129
; LENGTH: 1278
; TYPE: PRN
; ORGANISM: Homo sapiens
US-11-191-274A-129

Query Match 34.8%; Score 2402.5; DB 6; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;

Qy 7 RGLWALLRLAQSEPYTHIQPGYCAFDVDCG-----KNPELSSGLMTLSNVSCLS 59
Db 4 RGLAUGLLLLCPAQVFSQ-----SCVMYGEGLIAYGDKRYNCEYSG-----46
Qy 60 NTPARKITGDHLLILQKICPRLYTGNTQACCSAKQLVSLASLSITTKALLTRCPACSDN 119
Db 47 --PPKPLPKDGYDLVQELCPGFFFG-NVSLCCDVRQLQTLKDLNLQLPLQLSRCPSCFN 103
Qy 120 FVNLHCHNTCSNQSLFINVTR-----VAOLGAGQLPAVVAYRAFYQHSFABQSYDSCSRV 175
Db 104 LLNLFCELTCSRQSOFLNVTATEDYDVPVNTQTKNKELOYYVQCSFANAMYNACRDV 163
Qy 176 RYPAATLAVGTWCGYVGSALCNAQRWLNFGQDTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLGGKADA-CNATNWIEMFNKONGQAPFTTTPVFSDFPVH-----215
```

```
Qy 229 VSGSQPLNEGVARCNESQGDVATCSCQDCASCAIARPO-----ALDSTFYLG 279
Db 216 ---GKEPMNATKGCDESDEVTPACQDCSCIVCGPKQPQPPPPAPWITLGLDAMVIM 272
Qy 280 QWPGSLAVLIIILCSVPAVVTILL-----VGFVAPARDKSNWVDPKKGTSL 326
Db 273 WITTYMAFLVFPFAGFAVWCYKRYFVSYPIDSNIAFSV-NASDKGE-----ASCC 324
Qy 327 DKLSFSTHTLLQOFFCGMGTWASWPLTILVLSVIPVVALAAGLVTELTDTDPVELWSAP 386
Db 325 DPVSAAFEGCLARLPTRMGSPFCVPGNPGCVIFPSLVFITACSSGLVFRVTNPFVDMWSAP 384
Qy 387 NSQARSEKAFHDOHGFPPFRNQVILTAPNRSYRVSLLLPK-NFSGILDLDLLELL 445
Db 385 SSOARLEKEYFYQHGFPPFRTEQLIIRAPLTDKHIYQYPSPGADVPFGPGLDQILHOVL 444
Qy 446 ELQERLRHLQVMSPEAQRNISLQDICYAPLNPDNTSLYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASYNDETVTIQCICLAPLSPYNT---NCTILSVNLVYFQNSHSLDHKK 499
Qy 506 NOTLMGOTSQVDWKDHFLYCANAPLTFKDGTAALALSADYAPVFPFLAIGYKGDYS 565
Db 500 GDDFF---VVADYHTHFLYCVRAPASLNDTSLHDECLGTFGGVPFWLVGSDYQNYN 556
Qy 566 EBAALIMTSLNYPAGDPRLAOKLWBEAPFLEEMRAFORRMAGMPQVTTFAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAWKEFINFVKNYKN---PNLTISFTAERSIEDEL 613
Qy 626 NRTTAEDLPATSYIVIELYISLALGSSYSRVSVDKATILGCGVAVLVGAVMAAMG 685
Db 614 NRESDSDFVTWISYAIMFLYISLALGHYKSCRLLVDKSVLSGIAGIILVLSVACSLG 673
Qy 686 PFSYLGIRSSLVILQVPPVLVSVGADNIFIFVLEYQRLPRRPGEPREVIHGRALGRVAP 745
Db 674 VFSYIGLPLTLVIEVIFPLVLAGVDNIFILVQAYQORDERLQGETLDQOGLGRVLEVP 733
Qy 746 SMLLCSLSBAICFFLGAITPMPAVRTFALTGLAVILDLLQMSAFVALLSDSKRQAS 805
Db 734 SMFLSPSETVAFPLGALSVMFAVHTFSLFAGLAVFIDFLLOITCFVSLLLGLDIIKQEK 793
Qy 806 RLDVCCVKPQELPPPGQ-GEGLLGFQKAYAPFLLHMTIRGVLLFLALFGVSLYS 864
Db 794 RLDIFCCVRGAEDGTSVQASESCLFRFFKNSYSPLLKDKWMRPVIAIFVGLSPFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDYFLPLNRIFEVGAPVYFVTLGYNSSSEAGNAICSSAG 924
Db 854 NKVDIGLDQSLSPDDSYMDYFYSQYLHAGPPVYFVLEBEGHDYTSKQGNMVGCGMG 913
Qy 925 CNMFSTQKIQVATEPPEQSYLAIPASSWDDFIDWLT-SSCCRLYISGPNKDFCPS 983
Db 914 CNNDSLVQOIFNAQOLDNYTRIGFAPSSWIDDYFDWVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLKNCMSIT-MGSVRPSVEQFHKYLPWFILNDRPNIKCPKGLGAAYSYSVN--LTS 1040
Db 971 VDDPACVR-CRPLTPBEGKQRPQGGDFMRFLPMLSDNPNPKCGKGHAAYSSAVNILLGH 1029
Qy 1041 DGOVLASRFWAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFVPPYITNV 1100
Db 1030 GTRVGATYPMYTHVLTQTSADFDALKARLIASNT-ETMGINGS--AIRVFPYSVFTV 1086
Qy 1101 FYEQLTILPEGLFMLSCLVPTFAVSCLLGLDLSGLNLLSIYVILVDTVGFALWD 1160
Db 1087 FYEQLTIIIDTIFNLGVSIGAIPLVTVWLLGCELWSAVIMCATIAMVLNMFVWMLWG 1146
Qy 1161 ISYNAVSLINLVSAGMSVEFVSHITRSPAISTKTWLERAKEATISMGSAVPAVAMTN 1220
Db 1147 ISLNAVSLNVLNMSCGISVEFCSHITRAFTVSMKGSRAERAEALAHMSSVFSGITLTK 1206
Qy 1221 LPEGLVGLAKAQLOIIFPRLMLLTILGLLHGLVFLPVLVSYGPDVNP 1272
Db 1207 FGGIVVLAFAKSQIFQIFFRMYLAMVLLGATHGLIFLPVLLSYIGPSVKA 1258
```



```
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTTTPVSPDPVH----- 215
QY 229 VSGIOPLNEGVARNESGGDDVATCCODCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSCDCSVCKPQPPPPAPWPTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVTILL-----VGRVAPARDKSRWDPKGTSL 326
Db 273 WITYMAFLVFFGAFAVWCYKRYFVSYPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGOGGTWASWPLTILVLSPVVALAAGLVTELTTDPVLSWAP 386
Db 325 DPVSAAFEGCLRLRTRWGSFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
QY 387 NSQARSEKAFHDOHFGPPFRITNOVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDQHFQGFPRTEQLIIRAPLTDKHIYQYPVSGADVPFGPPLDIQLHQL 444
QY 446 ELQERLRHLQVWSPQAQRNLSIODICYAPLNPDTNLSYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASYDNETVTIQLDCLAPLSPYNT---NCTILSVLNYPQNSHVLHDKK 499
QY 506 NOTLMGOTSQVDKDHFLYCANAPLTFKDGTLALSCMADYCAPVFPFLAIGYKGYKYS 565
Db 500 GDDFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFCGVPFVPLVGLDYDQNTN 556
QY 566 EBAELIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMQVPTFTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAEKEFTNFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDELPIFATSYIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAVMAAG 685
Db 614 NRESDSVFTVVISVAIMFLYISLALGHKSCRRLLVDSKSLGIAGILVLSVACSIG 673
QY 686 FFSYLGISSVLQVWPLVSGADNIFIFVLEYQRLPRRGPPEVHIGRALGRVAP 745
Db 674 VFSYIGLPLTIVIEVPELVAVGVDNIFILVQAYORDERLOGETDQOLGRVLGEVAP 733
QY 746 SMLLCSLSBAICFFICALPMPAVRTFALTSLGLAVILDLLQWSAPVALLSDSKRQAS 805
Db 734 SMLFSFSEVTAFFGALSVMVAVHTFSLFAGLAVFIDFLQITCFVSLILGLDIKRQKN 793
QY 806 RLDDVCCVQBELPPPGQ--GEGLLGFFQKAYAPFLHWTIRGVVLLFLFALFGVSLYSM 864
Db 794 RLDDIFCCVRGABDGTSSQASESCLFRFKNSYSPILLKDWMPVITAI FVGVLFSIAVL 853
QY 865 CHISVLDELQALPKDSYLLDFLFLNRYFVGAPVYFTTILGYNFPSSAGHNAICSSAG 924
Db 854 NKVDIGLDQSLMPDDSYMVDYFKSISQYLHAGPPVYFVLEBHDYTSKQGNMVGGMG 913
QY 925 CNFSPQKIQVATEPPEQSYLAI PASSWDDPFDWLTLP--SSCCRLYISGPNKDFCPT 983
Db 914 CNDSLUQQLFNAQLDNTIRIGFAPSSWIDDYFDMVKFQSSCCRV---DNITDQFCNAS 970
QY 984 VNSLNLKNCMSGIT--MGSVRPSVEQPHKLPFLNDRPNIKCPKGLAAYTSVN--LTS 1040
Db 971 VVDPACVR--CRPLTPGKORPQGGPMRFLPFLNSDPNPNKCGKGHAAYSSAVNILLGH 1029
QY 1041 DQCVLASRPMAXHKPLKNSQDYTEALRAARELAANITADLRKVPGDPAFEVFPYITNV 1100
Db 1030 GTRVGATYFTVHTVLTQTSADFDALKKARLIASNTV--ETMGINGS--AYRVEPYSVFV 1086
QY 1101 FVEOYLITLPEGLPMLSLCLVPTFAVSCILGLDLSGLNLLSYMILAVDTVGPMALWD 1160
Db 1087 FVEOYLITLDDTIFNLGVSGLAIFLVTWVLLGCELWSAVIMCATIAMLVNMFGVWMLMG 1146
QY 1161 ISYNAVSLNLNLSAVGMSVEFVSHITRSPAISTKPTWLRARAEATISMGSAVFAGVAMTN 1220
Db 1147 ISLNAVSLNLNLSWMSGISEVFCSHITRAFTVSKGRSVERAEALAHMGSSVFSGITLTK 1206
QY 1221 LFGILVLGAKAQLIQIFFRNLNLLITLGLHLGLVFLVPLVLSYVGPVNPA 1272
Db 1207 FGGIVVLAFAKSQIQIFFRNVLAMVLLGATHGLIFLPLVLLSYIGSPSNKA 1258
```

RESULT 8

```
US-11-385-692-2401
; Sequence 2401, Application US/11385692
; GENERAL INFORMATION:
; APPLICANT: JOSELOFF, Elizabeth et al.
; TITLE OF INVENTION: KIDNEY DISEASE TARGETS AND USES THEREOF
; FILE REFERENCE: CL001591ORD
; CURRENT APPLICATION NUMBER: US/11/385,692
; CURRENT FILING DATE: 2006-03-22
; NUMBER OF SEQ ID NOS: 6044
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2401
; LENGTH: 1278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-385-692-2401
```

```
Query Match 34.8%; Score 2402.5; DB 7; Length 1278;
Best Local Similarity 39.8%; Pred. No. 1.5e-185;
Matches 522; Conservative 237; Mismatches 450; Indels 103; Gaps 25;
```

```
QY 7 RGMLLWALLRLAQSEPYTHIQPGYCAFYDECG-----KNPELSGSLMTLSNVCSLS 59
Db 4 RGLALGLLGLLLCPAQVFSQ-----SCVMYGCAGIAYGDKRYNCEYSG----- 46
QY 60 NTPARKITGDHLLILLOKICPRLYTGPTNTOACCSAKQLVSLASLSITKALLTRCPACSDN 119
Db 47 --PPRKLPGKGVDDLQVELCPGFPG--NVSLCCDVRQLTKDNLQLPLQLSRCPSCFYN 103
QY 120 FVNHLCHNTCSNQSLFINVTR----VAOLGAGQLPAVVAYEAFYQHSFAESYDSCSRV 175
Db 104 LLNLFCELTCSPRQSOFLNVTATEDYVDPTNQTNTNKLQYVYVQSQFANAMYNACRDV 163
QY 176 RVPAAATLAVGTMGVYGSALCNAORWLNFGQDGTGNGLAPLDT-----PHLEPGQA 228
Db 164 EAPSSNDKALGLLCKGDADA-CNATNWIEMFNKONGQAPFTTTPVSPDPVH----- 215
QY 229 VSGIOPLNEGVARNESGGDDVATCCODCAASCAPAIARPO-----ALDSTFYLG 279
Db 216 ---GMEPMNATKGCDESVDVETAPCSCDCSVCKPQPPPPAPWPTILGLDAMYIM 272
QY 280 QMPGSLVLIILCSFVAVTILL-----VGRVAPARDKSRWDPKGTSL 326
Db 273 WITYMAFLVFFGAFAVWCYKRYFVSYPIDSNIAFSV-NASDKGE-----ASCC 324
QY 327 DKLSFSTHTLLGQFGOGGTWASWPLTILVLSPVVALAAGLVTELTTDPVLSWAP 386
Db 325 DPVSAAFEGCLRLRTRWGSFCVRNPGCVIFPSLVFITACSSGLVFRVTTNPVDLWSAP 384
QY 387 NSQARSEKAFHDOHFGPPFRITNOVILTAPNRSYRYSLLGPK-NFSGILDLLLELL 445
Db 385 SSOARLEKEYFDQHFQGFPRTEQLIIRAPLTDKHIYQYPVSGADVPFGPPLDIQLHQL 444
QY 446 ELQERLRHLQVWSPQAQRNLSIODICYAPLNPDTNLSYDCCINSLLQYFQNNRTLLTLTA 505
Db 445 DLQIAIEN--ITASYDNETVTIQLDCLAPLSPYNT---NCTILSVLNYPQNSHVLHDKK 499
QY 506 NOTLMGOTSQVDKDHFLYCANAPLTFKDGTLALSCMADYCAPVFPFLAIGYKGYKYS 565
Db 500 GDDFF--VYADYHTHFLYCVRAPASLNDTSLHDPCLGTFCGVPFVPLVGLDYDQNTN 556
QY 566 EBAELIMTFSLNYPAGDPRLAQAKLWEAFLEEMRAFORRMAGMQVPTFTAERSLEDEI 625
Db 557 NATALVITPPVNNYNDTEKLQRAQAEKEFTNFVKYKN---PNLTISFTAERSIEDEL 613
QY 626 NRTTAEDELPIFATSYIVIFLYISLALGSYSWSRVMVDSKATLGLGGVAVVLGAVMAAG 685
Db 614 NRESDSVFTVVISVAIMFLYISLALGHKSCRRLLVDSKSLGIAGILVLSVACSIG 673
QY 686 FFSYLGISSVLQVWPLVSGADNIFIFVLEYQRLPRRGPPEVHIGRALGRVAP 745
```

```

Db 674 VPSYGLPLTLIVIEVPLVLAVGVNDIFILVQAVQORDERLOGETLDOQLGRVLGEVAP 733
Qy 746 SMLLCSLSEAICFFFLGALTPMAVTPALTSGVLAVILDFLOLQMSAFVALLSDSRQAS 805
Db 734 SMFLSFSSETAVFFLGALSVMPAVHTFSLFAGLAVFIDFLQITCFVSLGLGDIKROEKN 793
Qy 806 RLUDVCCVKPQLPPPGQ-GEGLLGGFFOKAVAPFLLHWITRGVLLFLFALFGVSLYSM 864
Db 794 RLIDFCCVRGADGTSVQASECLFRFFKNVSPILLKDWMPVIAIFVGVLSFIAVL 853
Qy 865 CHISVGLDQELALPKDSYLLDFLFLNRYFEVAGPVYFVTTILGYNFSSEAGMNAICSSAG 924
Db 854 NKVDIGLQSLMPDSDYMDYFKGISQYLHAGPPVYFVLERGHDTYTSKGQNMVCGMG 913
Qy 925 CKNFSFTQKIQVATEFPPEQSYLAI PASSWDDFIDMLTP-SSCRLYISGPNKDFCPST 983
Db 914 CNDSLVQOIFNAAGLDNYTRIGFAPSSWIDDYFDMVKPQSSCCRV---DNITDQFCNAS 970
Qy 984 VNSLNLCKNCSMIT-MGSVRPSVEQFHKYLPWFLNDRPNIKCPKGLAAYSTSVN--LTS 1040
Db 971 VDPACVR-CRPLTEGKQRPQGGDFMRPLFNLSDNPNKCGKGHAYSSAVNILLGH 1029
Qy 1041 DQOVLASRPMAYHKPLKNSQDYTEALRAARELAANITADLRKVPCTDPAFEPVYTIYNV 1100
Db 1030 GTRVGATVMTYHTVLQTSADFIDALKKARLIASNV-ETWINGS--AYRVFPYSVFV 1086
Qy 1101 FFEQVLTILPEGLFMLSCLVTPFAVSCLLIGLDRSLNLLSIVMILVDTVGFMAWD 1160
Db 1087 FFEQVLTIDDDTFIFNLGVSGLAIFLVTWVLLGCELWSAIVMCATIAMLVNMFVWMLWG 1146
Qy 1161 ISYNAVSLINLSAVGMSVEFVSHITRSFAITKPTWLERAKEATISMGSAFAGVAMTN 1220
Db 1147 ISLNAVSLVNUVMSGISEVEFCSHITRAFTVSMKSRVERAEALAHMGSSVFSGITIK 1206
Qy 1221 LFGIILVGLAKAQLIQIFFRNLMLITLGLLHGLVFLPVLISYVGPDPNPA 1272
Db 1207 FGGIVVLAPAKSQIQIFVFRMYLAMVLIGATHGLIFLPLVLSYIGPSVUKA 1258

RESULT 9
US-11-301-094-2
; Sequence 2, Application US/11301094
; GENERAL INFORMATION:
; APPLICANT: Levitan, Diane J
; APPLICANT: Smith, Marsha
; TITLE OF INVENTION: FUNCTIONAL ASSAYS FOR CHOLESTEROL ABSORPTION INHIBITORS
; FILE REFERENCE: JB06242US01
; CURRENT APPLICATION NUMBER: US/11/301,094
; CURRENT FILING DATE: 2005-12-12
; PRIOR APPLICATION NUMBER: 60/636,390
; PRIOR FILING DATE: 2004-12-15
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 2
; LENGTH: 1383
; TYPE: BPT
; ORGANISM: Caenorhabditis elegans
US-11-301-094-2

Query Match 19.4%; Score 1341.5; DB 6; Length 1383;
Best Local Similarity 27.5%; Pred. No. 2e-99;
Matches 369; Conservative 257; Mismatches 561; Indels 155; Gaps 40;

Qy 76 KICPLRYTGPNTQACSAKQVLSLEASLITKALLTRCPACSDNFVNLHCHNTCSPNQSL 135
Db 63 BECPHLLTGDN-KLCCTPSQABGLTKQIAQARHILGRCPSCFDNFAKLWCBEFTCSPNQOD 121
Qy 136 FENVTRVAOL--GAGOLPAVVAEAF----YQHS--FAEQSYSDCSRVRVPAATLAVG 186
Db 122 FVISEMKIEKGEFTPEYQAEAVVTVFRLSTDFAEGNFSCKDVTFGQPALRV- 180
Qy 187 TWCGVVGSAALCNAQRWLNFGQDGTGNGL-APLDITFHLLEPGQAVSGIOP-LNEGVARCN 244

```

```

Db 181 -WC---TSTPCTLTNWLEFIQTQNLDLNPIHTKFLLYDPIKTPPSDRSTYMNVNFTGCD 236
Qy 245 ESQDDVATCSQODCAASCPAIARQALDSTFYLGOM-----PGSLVLIILICSFAVVT 299
Db 237 KSARVGVMPACSTSEC--NKEYANLIDLDGKTSQTCNVHGIACINIFVMLAFIGSLAV 294
Qy 300 ILLVGF-----RVAPARDKSMVDPKKGTSLSDKLSFSFTHLLGQFF 341
Db 295 LLCVGFVTSDEDTNLRQTSQSEESPRNRK-----RTGANI-----HNFMENNA 342
Qy 342 QMGWTWASWBLTILVLSVIPVVALAAGLVFTELTDPELWSPNAPSQAQSKAFHQHF 401
Db 343 RDIGMMAGRNPKSHFFIGCAVLIFCLPGMIVHKESTNVVDMWSSPRRARQOEVMFNANF 402
Qy 402 GPPFRTNQVILTAPNRRSSRYRDSLLGPKNPSGILDLDLLELLELQELRLHLQWSPEA 461
Db 403 GRPQYQQIMLL--SHRDFQSSGKLYGP-----VFHKDIFELFDILNAIKNISTQDSG 455
Qy 462 ORNISLODICVAPLNPDNTSLYDCCINSLLQVFNQNRITLLILLTANOT-----L 509
Db 456 -RTITLDDVCYRPMGPG---YDCLINSPTNYFQGNKEHDMKSNKEETVSEDDADPYF 510
Qy 510 MGQTSQVDWKDHFLYCANAPLTFKDGITALASCMDYGAAPVFPFLAIGYKGYSEABA 569
Db 511 SSEATTDEMNHMAACIDQPMQK--TKSGLSCMGTYGPGSAPNM-VFGKNTNHOAANS 567
Qy 570 LIMFSLNNYPAGDPRLAQAKLWEEAFLEEMRAFORMMAGMFQVTFTAERSLEDEINRT 629
Db 568 IMMTILVTQ--RTEPEIQABELWEKEFLKCKBYREKSPKVI-FSFMABERSITDIBENDA 624
Qy 630 AEDLPFIPATSYIVIFLYISLALGSY-----SSWSRVMDVSKATLGLGVAVVGLVMAAM 684
Db 625 KDEIVTVVIALAFILGYVTFSLGRYFVCENQLMS-ILVHSRCLGLMLSVIINLLSFCWS 683
Qy 685 GFFSYLGRSSILVILQVVPFLVLSVGADNIFIFVLEY--QRLP---RRPGEPRVHIGRA 739
Db 684 GIFSMFGIHPVKNALVVOQFVVVTLIGVCRTFMVVKYQAQRVSMFMPSPDQCPEI-VGMV 742
Qy 740 LGRVAPSMLLCSLSEAICFFFLGALTPMAVTEALTSGVLAVILDFLOLQMSAFVALLSDS 799
Db 743 MAGTNPAMPFSSSLGCASFIFGGFTDLPAIRTFCLYAGLAVLIDVVLHCTTIFLALFVWD 802
Qy 800 KRQEAASRLDVCCKVPQELPPPGQEGEGL-----LLG-----FFQKAYAPFLLHWI 844
Db 803 QRELNG-----KP-EFFFPYQIKOLLGAYLIGRQATDFTMTQFFHFQVAPFLMHRM 853
Qy 845 TRGVVLLFLFALFGVSLYSMCHI SVGLDQELALPKDSYLLDFLFLNRYFEVAGPVYFT 904
Db 854 TRITIGIIFIASFITTVILSSKISVGFQDSMAFTEKSYISTHFRYLDKRFDPVGPVFTV 913
Qy 905 TLGYNFSSEAGMNAICSSAGCNPFSTQKIQVATEFPPEQSYLAI PASSWDDFIDWLT-P 963
Db 914 DGEIWDHPRDVQNKFCPTFGCSDTSFGNIMYAVGHTQTYLSGEMYNIDWILEMISRK 973
Qy 964 SSCRLYISGPNKDFCPSTVNSL-----NCLKNCM-----SITMGSV---RPSVEQPH 1009
Db 974 SPCCVKVYVHDN--TFCTNRNKSALDDDKACETCMDFDYVANSYPKSSIMYHRPSIEVY 1031
Qy 1010 KYLPWFLNDRPNIKCPKGLAAYSTSVNLTSQDQVLASRFMAYHKLPL--KNQDYTEAR 1067
Db 1032 RHLRHFLEDTNSECVFGGRASFKDAISFTSRGRIQASQFMTHFKKLSISNSDRIKAMD 1091
Qy 1068 ABELAANITADLRKVPCTDPAFEPVYTIYNVFEQVLTILPEGLFMLSCLVTPFAVS 1127
Db 1092 TARMVSRRLERSI-----DQAHVFAYSKIPFPPEYEQYSTIMPILTQTQFITVGVFGII 1145
Qy 1128 CULLGLDRSLNLLSIVMILVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITR 1187
Db 1146 CVTLGIDVKGACAVICQVSNFYHIVAFWYIFNIPNALSATNLVMSGILIEFSVNLK 1205
Qy 1188 SPAISTKPTWLERAKEATISMGSAFAGVAMTNLFGILVGLAKAQLIQIFFRNLMLIT 1247
Db 1206 GYACSLRQAKDRABSTVGSIGPIILSGFVVTMAGSTMFLSGAHLQIITVYFKPLITI 1265

```



```

; APPLICANT: Qin Zhang
; APPLICANT: Agnes Chopplin
; TITLE OF INVENTION: Targeted Gene Disruptions, Compositions
; TITLE OF INVENTION: and Methods Related Thereto
; FILE REFERENCE: R-DB-13
; CURRENT APPLICATION NUMBER: US/11/337,244
; CURRENT FILING DATE: 2006-01-20
; PRIOR APPLICATION NUMBER: US/10/669,143
; PRIOR FILING DATE: 2003-09-23
; PRIOR APPLICATION NUMBER: US 60/413,647
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/413,666
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/413,653
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/413,646
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/413,625
; PRIOR FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: US 60/413,639
; PRIOR FILING DATE: 2002-09-24
; NUMBER OF SEQ ID NOS: 237
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 149
; LENGTH: 1182
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-337-244-149

```

[illegible]

Qy	818	-----Lpp-----PG-----QGBG--	826
Db	594	AHLTATVQAFTHCEASSQHVVITILPQOAHLLSPASDPGLSGELYSPGSGTRDLLSQBEGTG	653
Qy	827	-----LLLGFFOKAVAPFLLHWITRGVLLLLFLAFLGVSLSMCHISVGLDQ	873
Db	654	QOACRPLLCAHWTLAHFARYQAPLLOTRAKALVLLFFGALLGLSLYGATILVQDGLAL	713
Qy	874	ELALPKDSYLLDYFLNLRYPEVCAPVYFVTTILGYNFS--SEAGMNAICSSAGCNPFSQTQ	932
Db	714	TDVPRGTKEHAFLSAQLRYSIL--YEVALYTOGGFDYAHQSRAI-----FDLHQ	761
Qy	933	KIQVATEFPQSYLAIPAS-----SWVDDF-----IDWLTPSSCCRLYIGPNK	976
Db	762	RFSSL-----KAVLPPPATQPRTWLHYRSYRSLQGIQAAPDQDWSAGRICTHSYRNGSED	816
Qy	977	D----KFCPSTVNSLCLKNCMSITM-----CSVRPSVEQFKYLPWFFLNDRPNIKCPK	1026
Db	817	GALAYKLLIQTGNAQEPDPSQUTTRKLVNKGELIPP--ELFYMGLTVWYSSOPL-----	869
Qy	1027	GGLA-----YST-----SVNLTSDGQVLASRFMAYHKPLKNSQDYTEALR	1067
Db	870	GLAASQANFYPPPPPEWLHKDYDTTGENLRIPAAQPLEFAQFFFLHLGLQKTADFVEAIE	928
Qy	1068	AARELANI--TADLRKVPGTDPAFEPVFPYITITNVFEQYILTILPEGLFMLSCL--LVPTF	1124
Db	929	GARAACTEAGQAGVHAYPSPGF-----LFWEQYLG--RRCFLLAVCILLVCTF	976
Qy	1125	AVSCLLGLGLDLRGLLNLISVILYDTVCFMALWIDISYNAVSLINILVSNAGVMSVEFVSH	1184
Db	977	LVCALLLLSPWTAGLI-VLVLAMWTVFELFGIMFLGKLSAIPVWILVASIGVGEFTVH	1035
Qy	1185	ITRSFAISTKPTWLERAKEATISGSAVFAGV--AMTNLPGLVLVLGLAKAQIOLIFFFR	1241
Db	1036	VALGFLTSHGSRNLRAA-----SALQGTAPVTDGAVSTILGLLLLAGSNFDFIIRYFV	1090
Qy	1242	LNLLITLLGLLHGLVFLPILSVXGP	1267
Db	1091	VLTVTLTLLGLLHGLLFPVLLSTLGP	1116

RESULT 15

```

US-60-772-265-1197
; Sequence 1197, Application US/60772265
; GENERAL INFORMATION:
; APPLICANT: BOUKHAROV, ANDREY A.
; APPLICANT: DU, ZIJIN
; APPLICANT: GUO, LIANG
; APPLICANT: HRESKO, MICHELLE C
; APPLICANT: KOVALIC, DAVID K
; APPLICANT: ZHAOLONG, LI
; APPLICANT: LU, MAOLONG
; APPLICANT: MCCARTER, JAMES P
; APPLICANT: MILLER, NANCY M
; APPLICANT: VAUDIN, MARK
; APPLICANT: WILLIAMS, DERYCK J
; APPLICANT: WU, WEI
; TITLE OF INVENTION: IDENTIFICATION AND USE
; OF INVENTION: FOR CONTROL OF PLANT
; FILE REFERENCE: MNDI-002USP1
; CURRENT APPLICATION NUMBER: US/60/772,265
; CURRENT FILING DATE: 2006-02-10
; NUMBER OF SEQ ID NOS: 1919
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1197
; LENGTH: 891
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: genomic DNA sequence=
US-60-772-265-1197

```

Query Match

7.0%: Score 482: DB 8: Length 891:


```
Db 997 IAAVLVCSGTYN--CDRLCKMKWNEV-----INP--EGFVNYLTGFNVD--NM 1044
QY 1023 KCPKGLAAYSTSVNLTSQGVLSRPMAYHKPLKNSQ-----DYTEALRAARELA 1073
Db 1045 MYVVSQASFPYPPGMEYNEKL--AKVDAEAPLLYSQMPFYQNDLIDTPAIVKMEIER 1102
QY 1074 ANITADLRKVPCTDPAFEVFPYTIITNVFVEQVLTILPEGLMLSLCLVPTPAVSCLLGL 1133
Db 1103 ATCE-----EYSEGLSNHPGIAFTFWEQVLT-LRNWLFQ-AICII-ALAVFCVISIL 1153
QY 1134 DLRSGLNLLSIVML--VDTVGFMALWDISYNAVSLINLSAVGMSVEFVSHITRSFAI 1191
Db 1154 MFNPAATLIMCIVITITIELGFGMLGMIKKNPISAVTLICAVGIGVEFTHAVELAP-L 1212
QY 1192 STKPTWLERAKGATISMGSAFAGVAMTNLPGILVLGLAKAQLIQIFFRMLNLLTLGL 1251
Db 1213 TALGTIDQRLSCLOHMFVYVYHGAISTFL-GVMMLVFSEFQVVTYFTYTTLLVALGV 1271
QY 1252 LHGLVFLPVILSVGPD-----VNPALEAQKEABEA 1283
Db 1272 FNLGCVLPVILTVGPKBELTDTGSSVLPPLPPPLRQQAES 1314

RESULT 17
US-11-360-355-133805
; Sequence 133805, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 133805
; LENGTH: 831
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (794)...(794)
; OTHER INFORMATION: xaa can be any naturally occurring amino acid
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=seqid 60228; Strand=-; Positions=1
; OTHER INFORMATION: -173,414-469,516-626,687-921,1040-1216,1282-1406,1470-1579,1627
; OTHER INFORMATION: -1785,1833-1933,1984-2145,2192-2241,2295-2457,2503-2695,2750
; OTHER INFORMATION: -2985,3164-3208
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_496761.1; Match level="QueryCoverage
; OTHER INFORMATION: =99%, HitCoverage=86%, E-value=0.0, Identity=48%"; Hit descriptio
; OTHER INFORMATION: =Patched Related (90.8 kD) (ptr-18) [Caenorhabditis elegans]
; OTHER INFORMATION: pir||T26683 hypothetical protein Y38F1A.3 - Cae
US-11-360-355-133805
Query Match 6.3%; Score 437; DB 7; Length 831;
Best Local Similarity 22.5%; Pred. No. 2.2e-26;
Matches 189; Conservative 139; Mismatches 303; Indels 210; Gaps 30;
QY 590 KLEWEAFLEEMAFQPRMAGMQVTFATERSLEDEINRTAEDLPITFATSYIVIFLYISL 649
Db 21 KKWEHSVYN--FAMNSRGDPLKLVHTSEGLVSEEVRTGIEVLPLMPLISLSVILFTVI 78
QY 650 A-----LGSY-----SSSRKRVMDVSKATLGLGGVAVVLGAVNMAAGFFSY 689
```

```
Db 79 TSLKRDQIRSKPWEALVGFCTVITSLKRDQIRSKPWEALVGFCTPILSVGASFTGLFW 138
QY 690 LGIR-----SSLVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPREVIHGRALGRV 743
Db 139 LGFEFLPGFBLPLVIVVPLIILAIGVDDVFIHLCHWAHTD--ESKDLRERVADLLGSA 196
QY 744 APSMLLCSLSBAICPFLGALTMPAPVTRTALTSLGLAVILDFLLQMSAFVALLSLDSKROE 803
Db 197 GPSVITISLTNWLSTIGIATPTPAIRTFCLFISVAVLYAYLYQLFFYTAVMVIGAQRRA 256
QY 804 ASRLDVCCCVK--POELPPPGQ-----GEGLLGFPFKQAYAPFLFLW 843
Db 257 DERNAYLCIKRVPKRSKPVATLKKTNLWRLGSTNLWRLGSDLV-----DRYVDFVMSW 311
QY 844 ITR-----GVLLLFAL--FGVSLYSVMCHISVGLDQELALPKDSYLLDYFLNLR--YFEV 896
Db 312 TARIILALSLIYWAFSAYGVAIKVAQIKVGLTSEKFLDSDSPLELVLRLQNNVIFKEG 371
QY 897 GAPVVFVTTLVGNFSSAGMNAICSSAG-CNNFSPTQIKIYATEPPEQSYLAIPASS--- 952
Db 372 QOMAVFVNSPG---DLRKPMAVFNSPGDLRKPSPAIPKIMRILEREQSGNSGVGSSTOM 428
QY 953 WYDDFIDLWTFSSCCRLYISGPNKDFCFSTVNSLNLCKNCHMSITMGSVR-----P 1003
Db 429 WLNTYLPFIG-----LQNRGSIINRGSIDPRYKYLYDFP 461
QY 1004 SVEQHKYLPWF-----LNDRNIIKCPKGLAAYSTSVNLTSQGVLSRPMAYHKP 1055
Db 462 SIPEYHRWSHFVSLGPKEDCLNERP-----SCINKFPDCLNERPSCINKFF-FSTG 511
QY 1056 LKNS--QDYTEALRAARELAA-----NITADLRKVPCTDPAFEVFPYTIITNVFVEQVLT 1107
Db 512 FQNAVMSDRVLVLRWQRLAADYSQMLT-----VYEDFSMWADQMYADQLLS 560
QY 1108 I----LPEGLFMLSCLVPTFAVSCLLGL-----DLRSGLLNLLSIVMLVDTVTGFMAL 1158
Db 561 IPSVTIQTVAFALLCMTFVAFALLCMTFVLVMTFPTSISTILPGTACVLSINLGVGLLFY 620
QY 1159 WDISYNAVSLINLSAVGMSVEFVSHI-----TRSFATISTPTWL----- 1198
Db 621 WSIDLPISMTTTLMAIGLSVDFVAHISFHYKYGMTTTLMAIGLSVDFVAHISFHYKYGE 680
QY 1199 -----ERAKEATISMGSAFAGVAMTNLPGILVLGLAKAQLIQIF----- 1238
Db 681 TEDSRERLRLHALSSIANPMLQ-AALSTVLSMLVLIHAYMVQVFKVVLVVLVGLVHG 739
QY 1239 ---FFRLNLLITLLGLHLGLVPLVILSYV-----GPDVNP-----ALALEQKRA 1280
Db 740 LVVFVKVVLVVLVGLVHGLVVLVPPVYAAIPPOKSGATAKPKVPLNRRRTAVTXQKRRR 799
QY 1281 E 1281
Db 800 E 800

RESULT 18
US-10-461-673-16750
; Sequence 16750, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
```

```
APPLICANT: Wang, Dunrui
APPLICANT: Yamazaki, Victoria
APPLICANT: Ujwal, Manuhea L.
APPLICANT: Ma, Yunging
APPLICANT: Chen, Rui-Hong
APPLICANT: Ghosh, Malabika
APPLICANT: Weng, Gezhi
APPLICANT: Haley-Vicente, Dana
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 823
CURRENT APPLICATION NUMBER: US/10/461,673
CURRENT FILING DATE: 2003-06-13
PRIOR APPLICATION NUMBER: PCT/US02/29964
PRIOR FILING DATE: 2002-09-19
PRIOR APPLICATION NUMBER: US 10/245,014
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/323,739
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: PCT/US02/29636
PRIOR FILING DATE: 2002-09-18
PRIOR APPLICATION NUMBER: US 10/245,817
PRIOR FILING DATE: 2002-09-16
PRIOR APPLICATION NUMBER: US 60/323,349
PRIOR FILING DATE: 2001-09-18
PRIOR APPLICATION NUMBER: PCT/US02/29001
PRIOR FILING DATE: 2002-09-13
PRIOR APPLICATION NUMBER: US 10/243,552
PRIOR FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: US 60/322,511
PRIOR FILING DATE: 2001-09-13
PRIOR APPLICATION NUMBER: PCT/US02/25485
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 17116
SOFTWARE: pt_FL_genes Version 6.0
SEQ ID NO 16750
LENGTH: 783
TYPE: PRT
ORGANISM: Homo sapiens
US-10-461-16750

Query Match 51.1%; Score 349; DB 6; Length 783;
Best Local Similarity 21.5%; Pred. No. 2.7e-19;
Matches 193; Conservative 160; Mismatches 335; Indels 208; Gaps 42;

QY 444 LLEQLERHLQWSPQARNISLQDICYPALNPNTSLYDCINSLLQYFQW----NRT 499
DB 10 ILKJHAAVTKIQV--PRPGNYTFAHICI--LNNDKTCIADDIVH-VLEELKARATNRT 64
QY 500 LLLLTANQTLMGQTSQVDWKDHPLYCANAPLT-FKDGTAALSCMADYGAPVFPFLAIGG 558
DB 65 NFAIT-----YPIHLKDGRA-----VYNGHQLGG 89
QY 559 --YKGD-YSEALIMTSLNNYPAGDPRLAQAKLWEAFLEEMRAPORMAGMPQVTF 615
DB 90 VTVHSKDRVKSAEAIQLTYLQSI NSLNDWAER--WESSFCDTVRLFQKSNKVKWQPY 147
QY 616 TAERSLEDEINTTAEDLIPATSIYIVFLYISLALGSSWSRVMWDSKATILGLGGVAV 675
DB 148 TSS-SLRFDQKTSRVSERYLVTSLIV--VTMAILCCS--MQDCVRSPKWLGLGLVT 201
QY 676 VLGVMAAGMFTSYLGIRSLVLTQVVPFLVLSVGADNFIIFVLEYORLPRRPGPREVH 735
DB 202 ISLATLTAGIINLTGCKYNSTFLG-VPFVMLGHGLYGTFFMLSSW----RTREDQHVK 256
QY 736 IGRALGRVAPSMLLCSLSA---ICFPLGA--LTPMPAVRTFALTSGLAVIDLFLQMSA 790
DB 257 -ERTRAVYADSMLSFLSTAMVLTFTGIGASPPTNIEAARIFCCNSCIAIFENYLVLSF 315
QY 791 FVALLSLDSKQEARSLDVCCVKP-----QELP-----PPQGE----- 825
DB 316 YGSSLVFTGYIENNYQHSIFCRKVPKPEALQKPAWYRFLLTARFSEDTAEGEEANTYBS 375
```

```
QY 826 GLLLGFFQKAYAPFLHLWITRGVLLFLALFGVSLYSMCHISVGLDQBLALPKDSYLLD 885
DB 376 HLLVCLKRYCYCDWITNTVVKPFWLFLYLIYSFALMGVLYQSEGSLSNIVATATOTIE 435
QY 886 YFLPLNRYPEVGAPVYFVTTLGYNSSSEAGMAICSSACGNPFSFTQKIQYATEPEQSY 945
DB 436 YTTAQOQYFNSYSPVI-----GFIYIESIEYWNYSVQEDV 470
QY 946 LAIPAS---SWDDFDLWLPSSCCRLVIS--GPNKDKFCPSVTNSLNCNKCMSITMGS 1000
DB 471 LEYTKGFVRISWFSYLVNLR-----KLVNVTGLPKCNTDMLRNSP--LK----- 514
QY 1001 VRPSVEQPHKYLFWFLNDRPNIKCPKGGIAAYSTSVNLTSDGQVLASRFMAYHKPLK-NS 1059
DB 515 -----APQFSHQEDIIPSK---KYNDEVD-----VVASRMFLVAKTMEINR 553
QY 1060 QDYTEALRAARELANITADLRKVGTDPAPEVFPYTTINVFYEQVLTILPEGLFMLSIC 1119
DB 554 EELYDLLETLRL--SVTSKVFIV-FNPSF-----VYMDRYASSL--GAPLHNSC 599
QY 1120 LVPTFAV--SCLLGLDLRSGLNL---LSIVMILVDTVGFMALWDISYNAVSLINLVA 1174
DB 600 ISALFLLFSAFLVA-----DSLINWITUTVSVFEGVIGFTMLWKVELDCISVLCIIVG 655
QY 1175 VGMSEVFVSHITRSPAIS---TKPTWLERAKBATISMSGSAVFAGVAMTN-----LPGILV 1226
DB 656 INVTIDNCAPMLSTFVLGKDFTRTKWKNALV-----HGVAILOSYLCYIVGLIP 706
QY 1227 IGLAKAQLIQIFPFLNLLITLLGLLHGLVFLPVILSYVGVDPVNPALALEQKRAE 1282
DB 707 LAAPFNSL--TCTLFRCLFLIAFVTFPHCFAILFVILTFLPPS-----KKRKEK 754

RESULT 19
US-11-360-355-141351
; Sequence 141351, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijong
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141351
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 67774; Strand=+; Position=1
; OTHER INFORMATION: -32,241-399,575-746,940-1023,1087-1203,3133-3154
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_999487.1; Match level="QueryCovera
; OTHER INFORMATION: =97%, HitCoverage=14%, E-value=2e-44, Identity=468"; Hit descrip
; OTHER INFORMATION: =Niemann-Pick C disease protein [Sus scrofa] gb|AAD47090.1| Niem
; OTHER INFORMATION: =Pick C disease protein [Sus scrofa] sp
US-11-360-355-141351

Query Match 4.7%; Score 325; DB 7; Length 204;
Best Local Similarity 39.6%; Pred. No. 3.1e-18;
Matches 76; Conservative 35; Mismatches 65; Indels 16; Gaps 3;

QY 585 RLQAQKWEAFLEEMRAPQRMAGMFQVTFTAERSLEDEINTTAEDLIPATSIYIVIF 644
```

```
Db 11 RNAAWLAWEQAFTSLKSYK---SQQFSYSFMAERSLEDEISRQSNADAFVFLSYTFMF 67
QY 645 LYSIALGYSYSSWSRVMVDSKATLGLGGVAVVLGVAAMAGFSYLGIRSSVLIVQVVPF 704
Db 68 IYVAIALGHMRRCRASLHSHSKFTLGTGCIIGCVLSVSSSIGLGFALFNMSATLIIIEVEPF 127
QY 705 LVLVSGADNIFLVELEYQRLPR---RP-----GPREVHIGRALGRVAPSMKLCS 751
Db 128 LVLALGVDNIFILVHSYQRLANDVDRPLPDRRLANDVDRPLPDRIAQCEDVIPSNMLSS 187
QY 752 LSEALCFFLGLAL 763
Db 188 LSECLCFSLGLL 199

RESULT 20
US-11-360-355-120785
; Sequence 120785, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 120785
; LENGTH: 575
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (48)..(48)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (62)..(62)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 47208; Strand=+; Position=1
; OTHER INFORMATION: -143,1368-1591,1911-2145,2447-2488,2669-2778,2839-2916,3181-3276,
; OTHER INFORMATION: 3451,3674-3825,3889-3953,4200-4265,4698-4936,5224-5367
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAE59824.1; Match level="QueryCoverage
; OTHER INFORMATION: =98%, HitCoverage=38%, E-value=1e-148, Identity=48%"; Hit descrip
; OTHER INFORMATION: Hypothetical protein CBG03294 [Caenorhabditis briggsae]
; FEATURE:
; OTHER INFORMATION: Essential gene: C elegans homolog=ZK675.1; Phenotype=Unclassified
US-11-360-355-120785

Query Match 3.8%; Score 261; DB 7; Length 575;
Best Local Similarity 20.4%; Pred. No. 2.3e-12;
Matches 127; Conservative 85; Mismatches 191; Indels 220; Gaps 24;

QY 791 FVALSLDSKQREASR--LDVCCCVK----- 814
Db 24 YFAFDLDRKRRKAGRRDLEFLCYXKTKIGSPFINNRPTAFKIFYQNNYELLFTIAK 83
QY 815 ---PQLPPPGQEGIL---LLGFFQKAVAPFLHWHITGVVLLFLALFGVSLYSMCHI 867
Db 84 ISVSQWATTPSAETVQNTLHGFLHIVIPLRRRAKFAILVTWVAUFAUGIYGLSRS 143
QY 868 SVGLDQELALPKDSYLLDYFLFLNRYFVGAPVYFVTTILGNFS----- 911
```

```
Db 144 TIGLESDVLDNT-----APAAFLKTRDKLFSYPMAVVRGANVDEFA 187
QY 912 -SEGAMNAICSSAGCNPNFTQKIQTATEFPEQSYLAIPASSWDDFDIML----- 961
Db 188 AQHQIDLLRRDIGLSPF---VVKLPDGPESRY-----WLQFRDWLRGMQQRDLV 236
QY 962 -----TP-----SSCRLYISGPNKDKPC- 980
Db 237 AKRDGLLDNFDGDTAKHSPDLQIAYSAAKHHSPDLQIAYSACS----GQNYD--CA 291
QY 981 -----PSTVNS---LNCL-----KNCMISITMGSVRPSVEOPHKYLPWFLNDR 1019
Db 292 RKAGTVRLIDESGINTGTFYNYLWGYEYQMFYSVQASFPYPLRKLQCGPPGMNFI 351
QY 1020 PNICPKGGLAAYSTSVNLTSDDQVLAASFMAFHYHKLKNSQDYTEALRAARELANIT-A 1078
Db 352 PPAPKP-----LFSRIPFYCDGLRDTPTIMRMIREIRAISENYSRA 392
QY 1079 DLKRVPGTDPAPFVPPYTTITNVVEOYL-----TILPEGLFMLS-CLVPTFA----- 1125
Db 393 GLPNYP-EGVAFTVAFV-----FWEQYLHLDYLLVLAIGIISVSVCVISTIIFNPWAPL 447
QY 1126 -VSCILLGLDLRSGLNLNLSIVMILVDTVGFMALWDISYNAVSLINLSVAGMSVEFVSH 1184
Db 448 WTCV-----VASMVELAGFMGLAGVKLNPISSVTLITAVGIGVEFTAH 492
QY 1185 ITRSFATSKPTWLERAKEATISMGSAVFAGVAMTNLPGLIIVGLAKAQLQIIPFRLNL 1244
Db 493 VALAF-LTSLGTREERMVNAQHMFPVVIHGG--NSTLLGIIMLAFSDFDVQVYFVVL 549
QY 1245 LITLLGLLHGLVFLPVILSYVGP 1267
Db 550 ALVTIGLFLGLAVLPVLLAYVGP 572

RESULT 21
US-10-461-673-16871
; Sequence 16871, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yuning
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Wang, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; CURRENT FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
```

```
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt FL_genes Version 6.0
; SEQ ID NO 16871
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-16871

Query Match      3.6%; Score 252; DB 6; Length 505;
Best Local Similarity 20.9%; Pred. No. 1e-11;
Matches 124; Conservative 105; Mismatches 219; Indels 146; Gaps 24;

QY 756 ICFFLGA--LTPMPAVRTALTSGLAVIDFLQMSAFVALLSLDSKROBASRLDVCCCV 813
Db 8 IFPGMGASPTNEAVKVCQNCVSIILNYPFSPFGSCLVPAQQLQNRYSIFCK 67

QY 814 KP-----QBLPPPGQGE-----GLLIGFFQKAYAPFLHLHWITRGVVL 850
Db 68 IPSAEYLDRKPVWFQTVMSDGHQOOSHETNPYQHFIQHFLREHYNEMITNIYVKPVV 127

QY 851 LLPLALFGVSLYSMCHISVCLDDELALPKDSYLLDYFLPLNRYFVGAPV---YFVTTLG 907
Db 128 ILYLIYASFSMCLQISDGANIINLLASDSPSVYAMVQOKYFSNYPVIGYVYEPL 187

QY 908 Y-NFSSEAGNNAICSSAGCNGNFSFTQKIQVATEFPFQSYLAIPASSWDDDFDMLTPSSC 966
Db 188 YNNSVQDDLRCLS-----GFT-----AVSWVEQYQFLKVSN- 221

QY 967 CRLYISGPNKPKCPSTVNSLNCMLKMSITMGSV--RPSVEQFHKYLPMFLNDRNPKC 1024
Db 222 ---VSANKSDP-----ISVLQSSFLKXPEFQHP-----RNDIIF 253

QY 1025 PKGLAAYSTSVNLTSDGQVLASRFMAYHKPLKNSQ-DYTEALRAARELAANITADLRKV 1083
Db 254 SKAG-----DESNIIASRLYLVARISDRDKQKEITEVLEKLRPL--SLSKSIRFI 300

QY 1084 PGTDPAFEVFPYITITNVFYEOY-----LTLPEGLFMLSCLVPTFAVSCLLGLDLRSG 1138
Db 301 V-FNPSF-----VFMHYSLSVTVPVLING-FGVLLVLLTLFVLHPLUG-----N 344

QY 1139 LLNLISVMILVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAIKPTWL 1198
Db 345 FWLILSVTSBELGLVGLMTLWVDMDCISILCIYTLNFAIDHCAPLLFTFVLATEHRT 404

QY 1199 ERAKEATISMGSAVFAGVAMTNLPGLILVLAKAQILQIFPFLNLITLL-----1249
Db 405 QCIKSLQDHGTAL-----LQNVTSFLI-----GLVPLLFVPSNLTFTLFKLLLTGGC 453

QY 1250 GLLHGLVFLPVILSYVGPDPVNPALAEQKRAE-----EAVAAMVASCNHPHSRV 1299
Db 454 TLLHCFVILPVLTFFPPP--SKKHKKKKRKRKEREIECIEIQBNPDHVTIV 505
```

RESULT 22

```
US-11-360-355-141367
; Sequence 141367, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
```

```
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETEROERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141367
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 67790; Strands=; Position=1
; OTHER INFORMATION: -258,397-491,542-698,907-1035,1085-1126,1171-1386,1440-1631,1681
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAE72420.1; Match level="QueryCoverage
; OTHER INFORMATION: =80%; HitCoverage=24%; E-value=1e-102, Identity=51%; Hit descri
; OTHER INFORMATION: Hypothetical protein CEG19582 [Caenorhabditis briggsae]
US-11-360-355-141367
```

Query Match 3.3%; Score 229.5; DB 7; Length 465;

Best Local Similarity 21.3%; Pred. No. 6.1e-10; Matches 91; Conservative 71; Mismatches 158; Indels 107; Gaps 14;

```
QY 940 FPEQSY--LAIPASSWDDDFIDWLTSPSSCCRLYISGPNKDKFCPSVNSLNCI-----KNC 993
Db 43 FDDDLKXMAADALARTDDFL--LARKLJC-----SAGQRFNCTEMIQRGA 86

QY 994 MSITMGSVRPSVEQFHKL-PMFLNDRPNIKCPKG-----GLAAYSTSVN 1037
Db 87 KLIEDGRINP--RGFTNYLTAMFYQDNMMYVSYQAYVYVQAAPFPPTPLPMHFSAADESUV 144

QY 1038 LTSQGVGLASRFMAYHKPLKNSQDYTEALRAARELAANITADLRKVPGTDPAFEVFPYTI 1097
Db 145 PPAD--PLLYSQIPFTYNGNLTDTQSIQVMIKEIRAICORYSAD-----GLPVPYPSGI 194

QY 1098 TNVFQEYQITLPEGLFMLSCLVPTFAVSCLLGLDLRSLIN-----1141
Db 195 PTFWEQYKLKLT---FYLVSAILIIGVAILGAVLVLSIIFNPWAAAMVAIIVTMFN 251

QY 1142 -----LLSIVMILVDVTGFMALWDISYNAVSLINLVSAGMSVEFVSHITRSPAIKSTK 1194
Db 252 PWAAAMVAIIVTMTVELAGFMGVGKKNPISAVTLITAVGIGVEFTAHVVLAPLTS-- 309

QY 1195 PFWLERAKEATISMGSAVPAGV-----AMTNLPGLILVLAKAQILQIFPFLNL 1244
Db 310 ----LGSRDERMVACLEHMFIPVEHMFIPVHGGSLTLGLVLMASFSEFPVVKYFFVWMT 366

QY 1245 LITLGLHGLVFLPVILSYVGP-----DVNPALAEQKRAEEAAVAVM 1288
Db 367 ALVIGLINGLALLPVLLSLIGPPCEITPVNGITPVNGSNLLPCPASERYRDESPNCFT 426

QY 1289 VASCNHP 1295
Db 427 TFSSTHH 433
```

RESULT 23

```
US-11-214-063A-1670
; Sequence 1670, Application US/11214063A
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
```

```

; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/11/214,063A
; CURRENT FILING DATE: 2005-08-30
; PRIOR APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1670
; LENGTH: 542
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-214-063A-1670

Query Match      3.3%; Score 225; DB 6; Length 542;
Best Local Similarity 20.4%; Pred. No. 1.8e-09;
Matches 120; Conservative 106; Mismatches 216; Indels 146; Gaps 26;

QY 744 APSMLLCSEA---ICFFLGA--LTPMPAVRTALTSLGLAVILDFLLQMSAFVALLSD 798
Db 23 ADSMLSFSLTTAMYLVTFGIGASPTNIEAARIFCCNSCIAIFFNLYLVSPGSSLVPT 82
QY 799 SKRQEARSLDVCCVKP-----QELP-----PPQGE-----GLLGFQ 833
Db 83 GVIENNYQHSIFCRKVPKPEALQEPAMYRFLLTARFSDTAEGEANTYESHLLVCF 142
QY 834 KAYAPFLHWITRGVLLFLALFGVLSYMSCHISVGLDOELAKPDVSLDLYFLNRY 893
Db 143 RYCDWITTYKPVFVLYLYISFALMGYLOVSEGLSNIVATATQITTYTAQKY 202
QY 894 FEVGAPVYVFTLLGVNFSSEAGNACSSAGCNCNFTQIOYATFEPPQSYLAIPAS-- 951
Db 203 FSNYSPI-----GFVYESIEYMNSTVQEDVLEVTGKGV 237
QY 952 --SWDDFDLMTSPSCCLYIS--GNKDKFCPESTVNSLNCNKCMSITGMSVRSVQF 1008
Db 238 RISWFESLYLNR-----KLVSTGLPKNFTDMLRNSF--LK----- 273
QY 1009 HKYLPWFLNDRPNIKPKGGLAAYSTVNLTSDGQVLASRFMAYHKPLK--NSQDYTEALR 1067
Db 274 ---APQSFHQEDIFSK---KYNDEVD-----VVASRMFLVAKTMTNREELYDLLE 320
QY 1068 ABRELAANTADLRKVGTDPAFEVPPYITITVFEQYITILPEGLFMLSCLCLVPTFAV- 1126
Db 321 TLRL--SVTSVKYKFTV-FNPSF-----VYMDRYASSL--GAPLHNSCISALFLIF 366
QY 1127 -SCLLGLDLRSGLNL--LSIVMLVDTVGFMALWDISYNVSLINLVSAGVMSVEFV 1182
Db 367 FRAFLVA---DSLINVITLTVSVFVGIFMTLWKVELDCISVLCILYGINITYDNC 422
QY 1183 SHITRSFAIS---TKPTMLERAKEATISGSAVFAGVAMTN-----LPCILVLGLAKAQL 1234
Db 423 APMLSTFVLGKDFTRTKVKNALV-----HGVALQSYLCYIVGLIPLAAVPSNL 473
QY 1235 IQIFPRMLNLLTLGLLHGLVFLPVLVSYPGVDNPNALAEQKRAE 1282
Db 474 -TCTLFRCLFLIAFVTFHCFAILPVILTFLPSS-----KKGKREK 513

RESULT 24
US-11-360-355-141365
; Sequence 141365, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
```

```

; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 141365
; LENGTH: 422
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 67788; Strand=+; Position=1
; OTHER INFORMATION: -132,192-287,332-472,523-811,902-1230,1410-1688
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP 494384.1; Match level="QueryCoverage
; OTHER INFORMATION: =58%, HitCoverage=20%, E-value=2e-85, Identity=54%"; Hit description
; OTHER INFORMATION: =patched family member (pcc-3) [Caenorhabditis elegans]
US-11-360-355-141365

Query Match      3.2%; Score 222.5; DB 7; Length 422;
Best Local Similarity 25.1%; Pred. No. 1.9e-09;
Matches 78; Conservative 61; Mismatches 117; Indels 55; Gaps 13;

QY 540 LSCNADYGAVPFPFLAIGYK---GKDYSBAEALIMTFSLNYPAGD--PRLAQAK--L 591
Db 1 MNCPEB-----LIVGIRRDAGGVQIRHAEAFQTVFLVAS--AGDVYQRLIRSKNHR 50
QY 592 WEEAFLEEMRAPQRMRMGMFQVTTAERSLEDEINRTTAEDL----PIFATS----- 639
Db 51 MEKFGVSTQAHAGAITAAWQNRFT--KSIYDHTLNKPAQGIRVVHPLASTSIQDMLBQF 108
QY 640 ---YIVIFL-YISLALGSYSWSRV-----WVDSKATILGLGVAVVLGAVMAAMGFF 687
Db 109 SEQFFVIFIGVLMII--YAGMSQVHWQWFSVKSSCLLAIIGLVITLASVAGLGLS 166
QY 688 SYLGRSSLVILQVVPFLVLSVGADNIPFVLEYQRLPRRPGEPREVIHGRALGRVAPSM 747
Db 167 TAMNHFNAATTQIVPFLTLGLIGIDDMPELLHNYNDVLE---AVRQKEVAVLLKETGMSV 223
QY 748 LLCLSSEACFGLGALTMPAVRTFALTSGLAVIDLFLQMSAFVALLSLDSKROEASRL 807
Db 224 LITSINNLAFTGCLIPALRSFCGQVAILLTSNVLCIILLFPATLALDURRKAHR 283
QY 808 DV-----CCC 812
Db 284 DMSFAALFCCC 294

RESULT 25
US-10-461-673-16889
; Sequence 16889, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Ausudi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yuning
; APPLICANT: Chen, Rui-Hong
```



```

; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 16889
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-461-673-16889

Query Match      3.2%; Score 219; DB 6; Length 1137;
Best Local Similarity 20.8%; Pred. No. 1.7e-08;
Matches 156; Conservative 98; Mismatches 304; Indels 192; Gaps 28;

QY 621 LEDEINRTABDL-----PIFATSYIVIFLYISALGSSYSWSRWMD 663
DB 198 LADNTSVTGMGLKQELLHFLVQDTVYPLA-----LVAIFPGMALYLSLEFLTMV- 252
QY 664 SKATGLGG--VAVVLGAVMAAGPFSVLGIRSSILVILQVVPFLVLSVGADNIRIFVLE 720
DB 253 ----LLGVLSLLVAFPLYQVAFRMAYFFPVNLAALL-----LSSVCANHTLIF-PD 300
QY 721 YORLPRR--PGBPREVHIGALGRVAPSMLLCSLSBAICFFLGALTPMPAVRTFALTSGL 778
DB 301 LWRLSKSQLPSCGLAQVRGTRMHFGYLLLVSLGTTSAAFYASYLSRLPAVRCLALFMCT 360
QY 779 AVILDELLQMSAFVALLSDSKRQEARLSDVCCVKPQBELPPPGQEG-----LLIGFPQ 833
DB 361 AVLVLHAL-----TLVWLPASAVLHERYLARGCARRAR---GRWEGSAPRLLIALHR 410
QY 834 KAYAPLLHWHITRGVLLFLALFGVLSYSMCHISV-----GLDQELALP- 878
DB 411 LRLG---LRRAAAGTSRLLFQRLPCGVIKFRIYICWFAALAAAGAYTAGVSPRLRPT 467
QY 879 -----KDSYLLDYF-----LFLNRYFEVGAPEYFVFTIGYN----- 909
DB 468 LPPPGQVPRPSHPFERFAEYRQLFLFEQLPQEGGHPVVLVWGVLPVDTGDPDLPDS 527
QY 910 -----FS-----SEAGMAICSSAGCNPFSTQKIQTAYTEFPEQSYLAIPASSWYD 955
DB 528 NSSLRVDPAFASGAPQARWLLALCHRA--RQSPFDTLQEGW-----PTLCFVE 575
QY 956 DFIDWLTSPSCCLYISGNKDKFCPSTVNSLNCNCKMSITWGSVRPSVEQPHKYLPH- 1014
DB 576 TLQRWMEPSFCARL---GP--DLCCG-----HSDFPWA 603

QY 1015 ----FLNDRNRIKCPKGLAAYSTSVNLTSDGVLASRFMAYHKPLKNSODYTEALRAARE 1071
DB 604 PQFFLHCLKWMALQEGPDGTQDLGLRFRDAHGS--LAALVLQFQTNFRNSPDYNQTLFYNE 662
QY 1072 LAANITADLRKVRGTDPAPEVEPPTITNTVFFYQYLITLPEGLPMLSLCLVPTFAVSCILL 1131
DB 663 VSHWLAARELGMAP---PGLRRGWFTRSLBLYSLQHSLSLSTEPVAVVLGALALAPAT--LLL 717
QY 1132 GL--DLRSGLLNLISIVMLVDTVGFMALWDISVNAVSLNLVSAVGMSEVYFVSHITRSPA 1190
DB 718 GTWNPVPLSLFSAVAVAGTVLLTVGLLVLEWQNTARALFLSNVSLGSLVDFVTNYCISYH 777
QY 1191 ISTKPTMLERAKEATISMGSVAVGAMTNLPGLVILGLAKAQLIQIPFRLMLLTLGLG 1250
DB 778 LCPHPDLRSVAPSLRQTSCTATVGAALFAAGVLMPLATV-----LLYKGLG 825
QY 1251 LHLGLV-----FLPVILSYVGPVYN 1270
DB 826 IILMMVKVSCGFASFFFQSLCCFFGPEKN 855

RESULT 26
US-10-461-673-10541
; Sequence 10541, Application US/10461673
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhi Wei
; APPLICANT: Yang, Yonghong
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Wang, Dunrui
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Ujwal, Manusha L.
; APPLICANT: Ma, Yunqing
; APPLICANT: Chen, Rui-Hong
; APPLICANT: Ghosh, Malabika
; APPLICANT: Weng, Gezhi
; APPLICANT: Haley-Vicente, Dana
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 823
; CURRENT APPLICATION NUMBER: US/10/461,673
; PRIOR FILING DATE: 2003-06-13
; PRIOR APPLICATION NUMBER: PCT/US02/29964
; PRIOR FILING DATE: 2002-09-19
; PRIOR APPLICATION NUMBER: US 10/245,014
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,739
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: US 10/245,817
; PRIOR FILING DATE: 2002-09-16
; PRIOR APPLICATION NUMBER: US 60/323,349
; PRIOR FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: PCT/US02/29001
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 10/243,552
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 60/322,511
; PRIOR FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: PCT/US02/25485
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 17116
```


Db 146 KF-----WLRDYQFVEQDPEAR--IKPENLDKYEIKFFIXSAGLVFRKLEPS 191
QY 998 MGSVRSVEQFHKYLPWFLNDRNINIKCPKGLAAYSTSNVLTSDQVLAIRFMAVHKPLK 1057
Db 192 MRN-----ELAEFLQW-----PEF---QGFLSANS-----ELVLSRFFF----- 223
QY 1058 NSQDYTEALR--AARELAANITADLRKVPCTDPAFEVFPVTITN---VFVEQVLTTLPEG 1112
Db 224 TTASYGTLEBNSKREL---LLQWRVADSVRSF-VFLFTVQEDDAKFLDIPMTIPQT 279
QY 1113 -----LPMLSLC-----LVPTFAVSCLLGLDLRSGLNLLSIWV--ILVDT 1152
Db 280 IQSSACTLLCMFLVCLLEMTNPGALLVSNFAISTCIGV---FGIQLSLGTLDDPIMSV 336
QY 1153 VGFMAWLDISYNAVINLVSAGMSVEFVSHITRSF-----AISTKPTWLERAKEAT 1205
Db 337 FGQSLGTLDDPIMSAIMSGFSDVIPAIAHYFYKTSVGEKAGTSLDVEQRLMDC 396
QY 1206 ISMSGSAVF-AGVA-MTNLPGLILVLGLAKAQLI-QIEFFRLNLLITLLGLLHGLVFLPVIL 1262
Db 397 ATIGFPVLEAGLSTLCTSSLPVDLHMARVFAITVFAITVVAIGLLHGLLVIPVNF 456
QY 1263 S 1263
Db 457 S 457

RESULT 31
US-11-360-355-140206
; Sequence 140206, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 140206
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 66629; Strand=-; Position=1
; OTHER INFORMATION: -151,199-261,383-478,553-695,766-905,1296-1406
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAE71975.1; Match level="QueryCoverage
; OTHER INFORMATION: =91%, HitCoverage=25%, E-value=3e-82, Identity=68%"; Hit descrip
; OTHER INFORMATION: =Hypothetical protein CBG19049 [Caenorhabditis briggsae]
US-11-360-355-140206

Query Match 2.0%; Score 135.5; DB 7; Length 252;
Best Local Similarity 24.2%; Pred. No. 0.01;
Matches 45; Conservative 48; Mismatches 56; Indels 37; Gaps 10;
QY 1099 NVFVEQYL-----TILPEG---LFMLSCLV-----PTEAVSCLLGLDLRSGLNLLS 1144
Db 66 NFFSQMLEKNTTLLSSGLTALLAMIFCVLIADSSIVFWFSMLVSMDI--GVCGYL- 122
QY 1145 IVMLVDTVGFMAWLDISYNAVINLVSAGMSVEFVSHITRSFPAISTKPTWLERAKE- 1203
Db 123 -VSMDIGVCGYLSLGSDDLPTTVVNLMSIGLCIDFATHGYR-----TYSRCRDP 174
QY 1204 ---ATISMS-----AVFAGVAVTNLPGLILVLGLAKAQLIQIEFFRLNLLITLLGLLHGLV 1256

Db 175 DQISDSLGAIGWPVVOAGV--STFLGIIVMLLVPSHVVRM--AQTMLVAVATGLFHGLF 230
QY 1257 FLPVIL 1262
Db 231 LUPILII 236
RESULT 32
US-10-724-972B-7050
; Sequence 7050, Application US/10724972B
; GENERAL INFORMATION:
; APPLICANT: DOUCETTE-STAMM, LYNN
; APPLICANT: BUSH, DAVID
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: 47040.0007US
; CURRENT APPLICATION NUMBER: US/10/724,972B
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; NUMBER OF SEQ ID NOS: 7546
; SOFTWARE: Patent in version 3.3
; SEQ ID NO 7050
; LENGTH: 808
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-724-972B-7050

Query Match 1.8%; Score 126; DB 6; Length 808;
Best Local Similarity 19.9%; Pred. No. 0.35;
Matches 105; Conservative 85; Mismatches 180; Indels 157; Gaps 25;
QY 841 LHWITR-----GVVLLFLALFGVS-----LYSM-----CHISVGLDQ 873
Db 69 LNMHPHGMFNLYVDGLGLLFSLLITGIGSLVLYSIGYLSKSEQLGNFYCYLLFMGA 128
QY 874 ELALP-KDSYLLDY-----FLENYFVEGAPVY-----FVTLGYNFSSEAGM 916
Db 129 MLGVVLSNFIILYFLWELTSFSPFLISFWREKASIIYGAQKSLITVIG-GLSMLGGI 187
QY 917 NAICSSAGCNFSTQKIQVATEFPPEOSYL-----AIPASSWVDDFDW 960
Db 188 --ILLSLATDTFSIQAMISKASDIQNSPPFIVLWILFMIGAFTKSAQVFFVWLPDAMEA 245
QY 961 LTPSSC-----CRLYISGPNKDKFCS-----TVNSLCLKNCMSITMGSVRPSVE 1006
Db 246 PTPVSAVYLHSAWVWAGLYLILARITPIFAISEGWVWTIT-----LVGLITLFWASLNATQ 301
QY 1007 QFHKYLPWFLNDRPNITKCPKGLAAAYSTSNVLTSDQVLAIRFMAVHKPLKNSQDYTEAL 1066
Db 302 HDLK-----GILAFSTVSQLGMINSLGIGAVSYHYQGANSQLYVAGF 344
QY 1067 RAARELAANI-----TADLRKVPCTDPAFEVFPVTITNVPFVEQYLT 1107
Db 345 VAAIFHLINHATFKGALFMITGGIDHSHGTGRDVKLGG---LLTIMPISFT-----LT 394
QY 1108 ILPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLLSIWMLVDTVGFMAWLDISYNAVS 1167
Db 395 VIT-----TSLMAGVPPF--NGFLSKKEFLSEMINVTHNLMSLNTLG-----ILLPIIA 442
QY 1168 LINLVSAGMSVEFVSHITRSFPAISTKPTWL--BRAKEATISMGSAVPAVMTNLPGLV 1226
Db 443 IIGSIFTFVYSIKFILHI---PFGSYKPEALPQAHESILM---LISPILTSL--VIV 494
QY 1227 LGLAKAQLIQIEFFRLNLLITLLG-----LLHGLVFLPVILSYVG 1266
Db 495 FGLFPSSILTSITTEPASVAVSQTNSNTABEPLHFGIT--PAFLSTIG 539

RESULT 33
US-11-360-355-149589
; Sequence 149589, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrew
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 149589
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 76012; Strand=-; Position=1
; OTHER INFORMATION: -51,383-466,817-1002,1773-1886,1992-2069
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=CAB66666.1; Match level="QueryCoverage
; OTHER INFORMATION: =89%, HitCoverage=16%, E-value=2e-48, Identity=63%"; Hit descrip
; OTHER INFORMATION: =Hypothetical protein CBG12004 [Caenorhabditis briggsae]
US-11-360-355-149589

Query Match 1.8%; Score 123.5; DB 7; Length 171;
Best Local Similarity 26.3%; Pred. No. 0.053;
Matches 41; Conservative 38; Mismatches 62; Indels 15; Gaps 7;
QY 614 TPTASRLDEINRTTAEDLPATSYIVIFLY-----ISLALGS-YSSWSRVMDSKAT 667
DB 12 TLYSQTLAEELKRNADTLVPFVFAFFILMFSLCMMSTCRGNYIDWTL-----SKPV 67
QY 668 IGLGVAVVLGAVMAAMGPFYSLGRSSVLQVFPFLVLSVGADNIFIFVLEYQRLPRR 727
DB 68 LAILGVANGMG-IVTAIGLLNCMSVPYN-DIVGVMPFLVAVGVDMFLMIAAVRTNR- 124
QY 728 PQEPREVIHGRALGRVAPSMML-CSLSEALCFELGA 762
DB 125 -ASPVCRMRGEAMADAISWITITIKTLITISFGGA 159

RESULT 34
US-10-536-606-20
; Sequence 20, Application US/10536606
; GENERAL INFORMATION:
; APPLICANT: Cosson et al.
; TITLE OF INVENTION: VIRULENCE GENES, PROTEINS, AND THEIR USE
; FILE REFERENCE: 25421-502NATL
; CURRENT APPLICATION NUMBER: US/10/536,606
; CURRENT FILING DATE: 2005-05-26
; PRIOR APPLICATION NUMBER: PCT/CH2003/00836
; PRIOR FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: 10/324,967
; PRIOR FILING DATE: 2002-12-19
; NUMBER OF SEQ ID NOS: 64
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20
; LENGTH: 1043
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-536-606-20
Query Match 1.8%; Score 121.5; DB 6; Length 1043;

Best Local Similarity 17.7%; Pred. No. 1.2;
Matches 236; Conservative 166; Mismatches 398; Indels 531; Gaps 58;
QY 134 SLPIINVTR---VAQLGAGOLPAV---VAYEAFYQHSFAEQSYDSRVRVPAATAAALVAG 186
DB 16 ALFISLAGLLVISKLPVAYQPNVAPPQITITATYP-----GASAKVLVD 59
QY 187 TMCVGYGSALCNAQRLNPFQ-DTGNGLAPLDITPHLEPG----- 226
DB 60 SVTSVLEESLNGAKGLLYPESSTNNSGTABIVVTP---EPGTDPLAQVDVQNRLLKAE 116
QY 227 ---QAV-GSGIQ-----PLNEGVARCNESQGDVATCSQDCAASCAPAIAR 268
DB 117 RMPQAVLTQGLQVQTSAGFLLIYALSKEGAQRSDTTALGDYAAARNINNELRLFGVGK 176
QY 269 PQALDSTFYLGQMPGSLVLIILCSVFAVVTILLGFRVAPARDKSKWDDPKK---GTS 324
DB 177 LQFFSSE-----AAMRV-----WIDPQKLVGFGLS 201
QY 325 LSDKLSFSTHTLLGQFFQGTWVASWPLTILVLSIPVVAL--AAGLVFTELTDTDPVEL 382
DB 202 IDD---VSNAIRGQNVQ-----VPAGAFGSAPGSSAQELTATLAVK 239
QY 383 WSPNSQARSEKAFHDQHFQPFRTNQVILTPAPNSSRYDSLLGPKNFSGLDLDLL 442
DB 240 GTLDDPQSGQVVLRLANEDGSLVRLADVARLELGKESYNISSRLNGTPTVGGAIQLSPGA 299
QY 443 ELLE---LQERLRLHQLVMSPEAQRNLSLODICYAPLNPDTNLSY-----DCINSLLQ- 492
DB 300 NAIQTATVLKQRLAELSAPFFE-----DMQYSV--PYDTSRFVDVAIEKVIHTLISA 349
QY 493 -----YFQNNRTLL--LTANQTLMGQTSQVMDKDHFLYCANAPLTFKDGFTALAL 540
DB 350 MYLVFLVMFLFQNVRYTLIPSIIVPVCLGLTWM-----YLLGFSVNMTMFGNVLA 404
QY 541 SCADYGAIPVPPFLAIGGKYGDYSEARALIMFSLNNYPAGDPRLAQAQKWEAPLEM 600
DB 405 GILVD-----DAIVVENVERIMA--EEGISA-----EATVKAM 437
QY 601 RAFORMMAGPQVTTAERSLEDEINRTTAEDLP--FATSYIVIFLYISLAL----- 651
DB 438 KQVSGAIVGITLV-----LSAVFLPLAFMHAGSVGVYQQFSVSLAVSILFS 483
QY 652 -----GSYSWSR-----VNVDSKATILGLGVAV 675
DB 484 GFALITFTFALCATLLKPIPEGHEKRGFGAFNRPARGVARTERYSLNSKLVARAGRFWL 543
QY 676 VLGAVMAAGPFYSLGRSSVLQVFPFLVLSVGADNIFIFVLEYQRLPRRPGPREVH 735
DB 544 VYAGLVAMLYGF-YLRLPEAFVPAEDLGMVVDV-----QLP--PGASR--- 584
QY 736 IGRALGRVAPSMMLCSLSEALCFELGALTMPAVTFA-----LTSGLAVILDLLQ 787
DB 585 -----VTDATGEBELRFLKSRREAVSFLIS 611
QY 788 MSAPVALLSLDSKROEASRLDVCCCVKQELPPQCGELGLLGFQKAVAPFLLHWITRG 847
DB 612 GFSP-----SQGDNAALAP-----PTFKDWSERG 636
QY 848 VLLFLFLALFGVLSYMSCHISVGLDOELALPKDSYLL----- 884
DB 637 -----AEQSAAAEIAALNEHFALPDGTVMVAVSPPPINGLNSGGFALRLMDR 684
QY 885 -----DYFL-----FLNRYFE--VCAPVYFV-----TTLYNYSSEAG 915
DB 685 SGVGREALQARDTLGLBIQTNPKEFLYAMMEGLAEAPQLRLIDREKARAFVSPETISG 744
QY 916 -MNAICSSAGCNPSPFTOKIOYATFEPPOSYLAI PASSWVDDFIDWLTSSCCRLYISGP 974
DB 745 TLSAAGFGEVINDFTNAGRQQRVVIOAEQG-----NRMTPESVLELYV--- 787
QY 975 NKDFPCPSTVNSLNCNKMCSITM--GSVRFSVEQFHKYPFWLNDNRPNIKCPKGLAAY 1032

Db 788 -----PNAAGNLVPLSAFVSVKWEGPV-----QLVRY-----NGYPSIRIVGDAAPGF 831
QY 1033 STSVNLTSGQVLASRFMAYHKLKNSQDYTEALRAARELANITADLRKVPGTDPAFEV 1092
Db 832 ST-----GEAMA-----EMERLASQLPAGI-----851
QY 1093 FPYTITNVFEOYLTI-LPEGLFMLSCLVPTFAVSCLLGLDLRSGLLNLASIVMIL-- 1149
Db 852 -GYEWGGLSYQEKVSAGQATSLFALAILVV-----FULLVALYESWIPLSVMLVPI 903
QY 1150 --VDTVGFMALMDISYNAVSLINLNSAVGMS-----VEFVSHITRSPAISTKPTWLE- 1199
Db 904 GAIGAVLAVWVGMSNDVFKVGLITIIGLSAKNAIILVEF-----AKELMEQG 952
QY 1200 -RAKEATISMGSAVAGVAMTNLP-----GILVL-----GLAKAQLIQIPFFRLNLITLL 1249
Db 953 HSLRDAAIIEARLRFRPIITSMNAFTLGIPLALASGAGASQRAIGTGVIGMLSATPL 1012
QY 1250 GLIHLGLVPLPV 1260
Db 1013 GVL-----FVPI 1019

RESULT 35
US-11-203-806A-11
; Sequence 11, Application US/11203806A
; GENERAL INFORMATION:
; APPLICANT: de Lange, Titia
; APPLICANT: Silverman, Joshua
; TITLE OF INVENTION: Nucleic and Amino Acid Sequences of
; FILE REFERENCE: Human Rf1 and Methods of Use Thereof
; FILE REFERENCE: 600-1-298N
; CURRENT APPLICATION NUMBER: US/11/203,806A
; CURRENT FILING DATE: 2005-08-15
; PRIOR APPLICATION NUMBER: 60/601,405
; PRIOR FILING DATE: 2004-08-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2426
; TYPE: PRT
; ORGANISM: Mus musculus
US-11-203-806A-11

Query Match 1.7%; Score 120.5; DB 6; Length 2426;
Best Local Similarity 19.9%; Pred. No. 5.2; Mismatches 106; Indels 341; Gaps 46;
Matches 184; Conservative 106; Mismatches 106; Indels 341; Gaps 46;

QY 11 LWALLRLAQSBPYTTIHQGYCAFYDECGKPNELSGSLMTLSNVSCLSNTP-----62
Db 259 LWPLFVKLLGK---TLHRSG-----SFINSLLQLEELGFRSGTPMIKKIAP 302
QY 63 ARKITGDHILLOKICPRLYTGNTQACSAKQVLSLASLITKALLTRCPACSDNPN 122
Db 303 AWKSLIDNPL-----NPDILCSAKRLKLLMQPLS-----S 333
QY 123 LCHNTCSNQSLFINVTVAOLGAGOLPAVVAEAFYQHSFAEQSDSCSRVVP-AAA 181
Db 334 IIVRTETLALTLEWVWVLLMLRGP-QLPA-----NFEQVCVPLIQS 374
QY 182 TLAVGTMCVYGVSALCNAQRWLNFQDGTGNGLAPLDITPHLLEPGQAVGSGIQPLNEGVA 241
Db 375 TISVDSIPSPQG-----NSSR-----GSASPGLSL-----LTPGH-----KGASPYGSPRG 415
QY 242 RCNESQDDVATCSQDCACSPATARPQALDSTFYLGOMPGSLVLIILCSVFVAVTIL 301
Db 416 NLSSTNGG-----NAATPEI-----QLLG-----LEMLLHF 441
QY 302 LVGFRVAPARDSKMV---DPKKGTSLSKLSFS---THLLGQFFQGWGTWVWSMPLTIL 356
Db 442 LLGPEVLSFAKHQKIVLSLEPLEHPLISPSFSSKYAHTL-----ITAV 485
QY 357 VLSVIPVVALAAGLVTELTUTDPVEL-WSAPNQSARSEKAFHQHGFPPFRTNQVILTAP 415

Db 486 HDSFVSVGKSDASDAVVSIAWKELISLVKSVTEAGNRKEK-----SGSEVLT- 532
QY 416 NRSSRYRDSLLIGPKNF--SGILDLDLLELELEQER--LRHLQWMSPEAQRNISLDICY 472
Db 533 -----LLKSLENIKSEVFPVSKTLVLMETVKGLPPKVLGSPAYQ--VANWDI-- 579
QY 473 APLNPD-----NTSLYDCCINSLLQFQNNRTLLLTANTQTLMGOTSDQDKOHF 522
Db 580 --LNGTPALFLTIQIFNNLLCEGVED--EKFLNLETIV---GCVLSGPTSPLAFSDSV 632
QY 523 LYCAN-----APLT-----FKDGTALALSCHADYGAPVFPPL 554
Db 633 LTVINQNAKQLWKEHLWRMSMIVSPDTVIHQNEVNOGDALHNFSALYGAULTPLIN 692
QY 555 AIGYKGYKDYGEAEALIMTFSLNYPAGDPRLAQAKWEAEFLBEMRAFORMMAGMFOVT 614
Db 693 HI-----FSAQTFTGTMK--ALLKTWSELY---RAFTR---CASIV 726
QY 615 FTAER-----SLEDEINRTTAEDLPATSPATSYIVIFLYISIALGSIY 654
Db 727 ATAENLCCBELSSKIMCSLEDEV---LSDLLFLDRISHIIVMVDCIDFSPYNNKYQPK 783
QY 655 -----SSWSRVNVDKATGLGSAVVLGAVMAAMGFFSYLGIRSSSLVILQVVPFLVL 707
Db 784 IKSPQRSDNR-----KKKEPLKGLASLPKLVKVIDTFTHTLSLK-----ETPSDTLL 832
QY 708 SVGADNIFIFVLEYQRLPRRPGEPREHVIGRALGRVA-PSML-----LCSISEAICFFLG 761
Db 833 AIG--NSIISML-----SNVFGHISLPSMIREIPATFTPLALLYENS 873
QY 762 ALTPMPAVRTALTSLAVIL-----DFLQMSAPVALLSLDSKRO-- 802
Db 874 KLDEAPKYVT--SLNNKLEKLGLEIVACLFQSYLGAYDSELEHLSPLLCVIFLHKNKQIR 932
QY 803 -----EASRLDVCCCVKPOELPP 820
Db 933 QOSALLWNATPAKATALVVPPELXP 957

RESULT 36
US-11-214-063A-2044
; Sequence 2044, Application US/11214063A
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/11/214,063A
; CURRENT FILING DATE: 2005-08-30
; PRIOR APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2044
; LENGTH: 768
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-214-063A-2044

Query Match 1.7%; Score 118.5; DB 6; Length 768;
Best Local Similarity 25.9%; Pred. No. 1.3;
Matches 51; Conservative 36; Mismatches 87; Indels 23; Gaps 8;
QY 633 LPIFATSYIVIFLYISIALGSSYWSRWVDSKATLGG--VAVVLGAVMAAMGFFSYLG 691
Db 3 LAFISSCIAALVYILTSCSVLSFFGI-----ASIGLSCLVALFLYHVWFGI---QYL 54

;/ TITLE OF INVENTION: 95431, 22245, 2387, 16658, 55054, 16314, 1613, 1675, 9569 OR
;/ FILE REFERENCE: 13424 MOLECULES
;/ CURRENT APPLICATION NUMBER: US/11/312,958
;/ PRIOR FILING DATE: 2005-12-20
;/ PRIOR FILING DATE: 2003-02-19
;/ PRIOR APPLICATION NUMBER: US 60/360,495
;/ PRIOR FILING DATE: 2002-02-28
;/ PRIOR APPLICATION NUMBER: US 60/370,121
;/ PRIOR FILING DATE: 2002-04-04
;/ PRIOR APPLICATION NUMBER: US 60/373,010
;/ PRIOR FILING DATE: 2002-04-16
;/ PRIOR APPLICATION NUMBER: US 60/373,908
;/ PRIOR FILING DATE: 2002-04-19
;/ PRIOR APPLICATION NUMBER: US 60/377,717
;/ PRIOR FILING DATE: 2002-05-03
;/ PRIOR APPLICATION NUMBER: US 60/379,949
;/ PRIOR FILING DATE: 2002-05-13
;/ PRIOR APPLICATION NUMBER: US 60/382,409
;/ PRIOR FILING DATE: 2002-05-21
;/ PRIOR APPLICATION NUMBER: US 60/385,280
;/ PRIOR FILING DATE: 2002-06-03
;/ PRIOR APPLICATION NUMBER: US 60/386,879
;/ PRIOR FILING DATE: 2002-06-06
;/ Remaining Prior Application data removed - See File Wrapper or PALM.
;/ NUMBER OF SEQ ID NOS: 64
;/ SOFTWARE: FastSeq for Windows Version 4.0
;/ SEQ ID NO 46
;/ LENGTH: 619
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
US-11-312-958-46

Query Match 1.7%; Score 116.5; DB 6; Length 619;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 107; Conservative 69; Mismatches 164; Indels 201; Gaps 26;

QY 536 TALALSCWADYCAPV-----PFPLAIGGYKGDYSEAEALIMTFSIN---NYPAGDP 584
DB 77 SVLAGLCAEFGARVPRSGSAYLSYVTVG-----ELWAFITGNLILSVIGTA 126

QY 585 RLAAQKLMEEAFLEEMRAFORRMAGMFQVTFATERSLEDEINRTTAAEDLIPATSYIVIF 644
DB 127 SVARA--WSSAF-----DNLIGN-HISKTLQCSIALHVPVLAE-YPDFALGLVLL 174

QY 645 LYSIALGSYSS-----WSRVMD----- 663
DB 175 LTGLLALGASESALVTKVFTGVNLLVLGFMISGFVKGDVHWNKLTEDYELAMAEALNDT 234

QY 664 -SKATILGLGV-----AVLGAVMAMGFFSYLG-----IRSLVIL 699
DB 235 YSLGLPGSGGFVFFGEGILRG---AATCFYAPVGFDCIATTGEBQNPQRSIPMGIVIS 291

QY 700 QVVPFLVLSVGDADNIFIFVLEYQRLPRRGPGEVHIGRALGRVAFPM-LIGSLSEACFCF 758
DB 292 LSVCFLAYFAVSALTLMPPYQLQPESLPEAFLYIGWAPARYVAVVGSLSALSTSL-- 349

QY 759 FLGALTMPAV-----RTFA-----LTSG-LAVILODFLLQMSAFVA 793
DB 350 -LGSMPFPRVITYAMAEDGLLPRVLARIHTGTRTPIATVVSGLIAAFMAFLKLTDLVD 408

QY 794 LLSLDSKRQEAASRLDVCCCV---KQGE----- 817
DB 409 LMSIGT-LLAYSLSVSCVILRYQDPQETKTGEVELOEAEITTESEKLTLMGLFPPLNS 467

QY 818 LPPPGQGE-----GLLIGFFQKAYAPFLHWM-----ITRGVVLILFLALFG--VSL 861
DB 468 IPTPLSGQIVVYCSSLAVLLTALCLVAQMSVPLLSGDLTWTAVVLLLIIGIIVI 527

QY 862 YSMCHISVGLDQEL-----ALPKDSYLLDYFLFLNR-----YFEVGAPVYFTYILG 907
DB 528 WRQOSSTPLFKVPALPLLPLMSIFVNITLYLMQMTAGTWARFGVMMLIGAIYF---G 583

QY 908 Y 908
DB 584 Y 584

RESULT 40

US-10-461-673-12167
;/ Sequence 12167, Application US/10461673
;/ GENERAL INFORMATION:
;/ APPLICANT: Tang, Y. Tom
;/ APPLICANT: Liu, Chenghua
;/ APPLICANT: Zhou, Ping
;/ APPLICANT: Asundi, Vinod
;/ APPLICANT: Wang, Jian-Rui
;/ APPLICANT: Ren, Feiyan
;/ APPLICANT: Zhang, Jie
;/ APPLICANT: Zhao, Qing A.
;/ APPLICANT: Xue, Aidong
;/ APPLICANT: Wehrman, Tom
;/ APPLICANT: Wang, Zhi Wei
;/ APPLICANT: Yang, Yonghong
;/ APPLICANT: Goodrich, Ryle W.
;/ APPLICANT: Wang, Dunrui
;/ APPLICANT: Yamazaki, Victoria
;/ APPLICANT: Ujwal, Manusha L.
;/ APPLICANT: Ma, Yuning
;/ APPLICANT: Chen, Rui-Hong
;/ APPLICANT: Weng, Gezhi
;/ APPLICANT: Haley-Vicente, Dana
;/ APPLICANT: Dmanac, Radoje T.
;/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;/ FILE REFERENCE: 823
;/ CURRENT APPLICATION NUMBER: US/10/461,673
;/ CURRENT FILING DATE: 2003-06-13
;/ PRIOR APPLICATION NUMBER: PCT/US02/29964
;/ PRIOR FILING DATE: 2002-09-19
;/ PRIOR APPLICATION NUMBER: US 10/245,014
;/ PRIOR FILING DATE: 2002-09-16
;/ PRIOR APPLICATION NUMBER: US 60/323,739
;/ PRIOR FILING DATE: 2001-09-19
;/ PRIOR APPLICATION NUMBER: PCT/US02/29636
;/ PRIOR FILING DATE: 2002-09-18
;/ PRIOR APPLICATION NUMBER: US 10/245,817
;/ PRIOR FILING DATE: 2002-09-16
;/ PRIOR APPLICATION NUMBER: US 60/323,349
;/ PRIOR FILING DATE: 2001-09-18
;/ PRIOR APPLICATION NUMBER: PCT/US02/29001
;/ PRIOR FILING DATE: 2002-09-13
;/ PRIOR APPLICATION NUMBER: US 10/243,552
;/ PRIOR FILING DATE: 2002-09-12
;/ PRIOR APPLICATION NUMBER: US 60/322,511
;/ PRIOR FILING DATE: 2001-09-13
;/ PRIOR APPLICATION NUMBER: PCT/US02/25485
;/ Remaining Prior Application data removed - See File Wrapper or PALM.
;/ NUMBER OF SEQ ID NOS: 17116
;/ SOFTWARE: PCT_FL_genes Version 6.0
;/ SEQ ID NO 12167
;/ LENGTH: 619
;/ TYPE: PRT
;/ ORGANISM: Homo sapiens
US-10-461-673-12167

Query Match 1.7%; Score 116.5; DB 6; Length 619;
Best Local Similarity 19.8%; Pred. No. 1.4;
Matches 107; Conservative 69; Mismatches 164; Indels 201; Gaps 26;

QY 536 TALALSCWADYCAPV-----PFPLAIGGYKGDYSEAEALIMTFSIN---NYPAGDP 584
DB 77 SVLAGLCAEFGARVPRSGSAYLSYVTVG-----ELWAFITGNLILSVIGTA 126


```
Qy 405 PRTN-----QVILAPNRSSRYSDSLLLPKPNFSGILDLDLLELLEQLERLHLYQWS 458
Db 364 YQNCIEPTSKLIDFPPIFY-----DVLYSTENQHOYLAV-- 403
Qy 459 PEAQNRISLQDICYAPLNPNTSLDYCCINSLLQVFQNNRTLLLTANTQTLMGQTSQVDM 518
Db 404 --PVLNLNLQH-----NKFVFNQDSNGSKWLLTRRFLV--DAVSG 440
Qy 519 KDHFYCANAPITFDKGTALALSCMA---DYCAPVFPFLAIGGYKGDYSEAE-----A 569
Db 441 RENDL--GTQPRVIRVATQISVHLVPNTINGNIYPLITITAYSDIDIKDANSQSVKVS 498
Qy 570 LIMTSLNNYPAG--DPLAQAKLWEEAFLEEM--RAFORKMAGMFQVFTFAERSLEDEIN 626
Db 499 FSVTYEMDGEAHVQTDIALGVGLAVLASLLKTAGWKRRIGSPMIDL----- 547
Qy 627 RTAEDLPFATSYIVIFLYISLALGSSYSSSRVMDVSKATLGLGVAVVLGAVNAAMGF 686
Db 548 QTVKRFVYAGDLANVFPIITVGTGLY--W---LIFFKAQ--KSVSVLLPMPVQEEERF 599
Qy 687 FSYLGIRSLVILQVFPVLVSVGADNIFIFVLEYQRLPRRPGEPREVHIGALGRVAPS 746
Db 600 VTYVGAFALKALQFLHLKLSQITID---VFFIDWER-----BK-----GKVLKAV--- 642
Qy 747 MLLCSLSAICPFGLATPMPAVRTF-----ALTSGLAVILDFLLOMSA 790
Db 643 -----EGEGVRSATVPVSTWRYTFVANENEIQTWRKKNLSLQVLTVL--FLEVVG 693
Qy 791 FVALLSLDS-----KQBEASRLDVCCVQKPELPPPGQEGLLLPFKQKAYAPFLHLWTR 846
Db 694 FKNALMDSSSLSRNPPIYAPYSCI-----LRYAVSAALWLAI 733
Qy 847 GVVLILLFALF-----GVSLYSMCHISVGLDQELALPKDSYLLDPLFLNRYFE 895
Db 734 GIQVVFVAVFYERFIEDKIROFVLCSSMSNISV-----FLLSHKCF--GYI 779
Qy 896 VCAPVY-----PVTLTGVNFSEAGMNAICSSAGC--NNFSPYKIQVATPEQSYLAI 948
Db 780 HGRSVGHADTNMEENWMLKREA--ENLCSQRLGVPTDGTQTFIAISNQM--RQHYDRI 836
Qy 949 PASSWVDDPIDLTPSSCRLYISGNKDKFCPTVNSLNCNKCMISITMGSRVPSVEQF 1008
Db 837 HET-----LIRKNGPARLLSS-----ASTFEQSIKAYHMKNFGLGSF----IDHV 878
Qy 1009 HKYLPWFLNDRNPKCPKGLAAYSTSVNLTSDDQVILASRFMAYHKPLKNSQDYTE 1064
Db 879 HKEMDYFIKDKLLE-----RILGMEFM---EPMEKSIIFYND 912
```

RESULT 43

```
US-10-703-799B-226
; Sequence 226, Application US/10703799B
; GENERAL INFORMATION:
; APPLICANT: Pompejue, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Habernauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; RESISTANCE AND TOLERANCE PROTEINS
; FILE REFERENCE: BGI-124CPN
; CURRENT APPLICATION NUMBER: US/10703,799B
; CURRENT FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: 09/603,208
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27

Query Match 1.6%; Score 112.5; DB 6; Length 697;
Best Local Similarity 18.8%; Fred. No. 3.5;
Matches 183; Conservative 111; Mismatches 311; Indels 369; Gaps 44;

Qy 341 FQMGWTWVASWPLTILVLSVIPVVALAAGL-----VFTELTTDPV--ELWSAPNSOARSEK 394
Db 2 FSKGHGFAYRF-----RRIVPLVVIATLALFVIFGTLGDRMSQEGWDDPGSSSTAAA 55
Qy 395 AFHQHQFGPPFRITNQ-----VILTAPNRSSRYSDSLLLPKPNFSGILDLDLLELLEQER 450
Db 56 RIELETFG--RDNDGDVLLFTAPEGTSFDAAEVF--SSISGYLD-----GLIE--- 100
Qy 451 LRHLQWSPQAQRNLSQDICYAPLNPNTSLDYCCINSLLQVFQNNRTLLLTANTQTLM 510
Db 101 -----NNPDEVH--HINS--YFD-----TRNQILL 121
Qy 511 QOTQOVMDKDFLYCANAPLTFKGTALALSCMADYGAPEVFPFLAIGGYK-----KDY 564
Db 122 S-----KGTQ-----TFAAL-GLKGDGEQTLKDF 145
Qy 565 SEABALIMTFSLNYPAGDPLRLAQAKLWEEAFLEEMRAFQRMRMAGMFQVFTFAERSLEDE 624
Db 146 REIEDQLHPDNL--AG-----GVTTTEVAGATAVADALDEGMAGD 182
Qy 625 INRTAEDLPFATSYIVIFLYISLALGSSYSSSRVMDVSKATLGLGVAVVLGAVNAAM 684
Db 183 ISRAEVFALPFVAILLIVF-----GS-----VVAAAMPLIV 214
Qy 685 GFFSYLGIRSSLIVIL-----QVVPFLVLSVGADNIFIFVLEYQRLPRRPGEPRE 733
Db 215 GILSLTGLSLGILAILAGFPQVNVFAQSVVTLGLGLAID-YGLFWVSRFREEMDKGTPVE 273
Qy 734 VHIGRALGRVAPSMLLCSLSSEACFFLGLALTPMPAVRTFALTGSLAVILDFLLQMSAFVA 793
Db 274 QAVATTTATAGKTIVFESAAMVAVALSGLFVFPQAFKSAVFAFCAISAVGLAALMSVTVLPS 333
Qy 794 LLSLDSKR-----QEASRLDVCCVQKPELPPPGQEGLLLPFKQKAYAFLLHW 843
Db 334 LFSMLGRKNIDKWSLRRTARTARLLE-----TIWTRVPAWAWRHAKA 375
Qy 844 ITRGVVLLPLALFGVSLYSMCHISVGLDQELALPKDSYLLDYLFLNRYFEVGPVFPV 903
Db 376 VTVGVVLLL-LAL-----TVPLTGKFKGGINETYLPP----- 406
Qy 904 TTLGYNFSSEAGMNAICSSAGCNPNFSFTQKTYATEPPEQSYLAIPASSWVDDFDLWLP 963
Db 407 -----ANDTRVAQE-----RPDE---APPAFR----- 425
Qy 964 SSCCELXYISGNKDKFCPTVNSLNCNKCMISITMGSRVPSVEQHFHYLPWFLNDRPNIK 1023
Db 426 TEPVKLVVVTGADNNQLIDIYVQA-----NEVE----- 452
Qy 1024 CPKGLAAYSTSVNLTSDDQVILASRFMAYHKPLKNSQDYTEALRAARELANITADLR-- 1081
```

```
Db 453 -----GLTDRFTAGATDDGTTVL-----TGIQD-----RSINEQVVEQLRAI 491
QY 1082 KVP-----GTDPAFEV-----FPYITINVEYOYLTLPEGLFMLSICLVPTFAVSC 1128
Db 492 SVEGEVEIGGTPAMEIESIEALFEKLLWALYIVLATFILMALVFGSV-TLPAKAIIM 550
QY 1129 LLLGLDLRSLNLSIVMLVDTVGFMALWDISYNAVSLIN-----LVSAN--GMSVEFV 1182
Db 551 TILMGATGILITLM-----FVDGVGASAL-----NFSGPLMSPVLIMAIYGLSTDY- 601
QY 1183 SHITSPALSTPTMLERAKEATISMGSAVFAGVAMTNLPGILVLGLAKAQLIQIFPFL 1242
Db 602 -----EVFLVSR-----MVEARDKGESTDRAIRYGAHT---GSIITAAALIMIVVCGAGF 650
QY 1243 NLLITLLGLLHGLIV 1256
Db 651 SEIVMMKYIAFGMI 664

RESULT 44
US-11-360-355-152528
; Sequence 152528, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 152528
; LENGTH: 201
; TYPE: PRF
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID_78951; Strand=+; Position=1
; OTHER INFORMATION: -147,625-754,987-1185,1232-1296
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit ID=NP_495662.1; Match level="QueryCoverage=79%, HitCoverage=13%, E-value=2e-46, Identity=57%", Hit description="patched family member (ptc-1) [Caenorhabditis elegans]"
; OTHER INFORMATION: emb|CAA86843.1| Hypothetical protein ZK675.1 [Caen
US-11-360-355-152528

Query Match 1.6%; Score 111; DB 7; Length 201;
Best Local Similarity 27.9%; Pred. No. 0.7;
Matches 34; Conservative 19; Mismatches 65; Indels 4; Gaps 2;

QY 592 WEEAFLEEM--RAFOR--MAGFQVTTAERLSLEDEINRTAEDLPATSYIVFLYI 647
Db 53 WARAFTDSLNYHAFNRDRDIEQSAVVHPLASTSISDMLAEFCDFNVAITLAGYLLMLYA 112
QY 648 SLALGSYSSWSRMVDSKATLGLGGVAVVLGVAMAGFFSVLGRSSVLIVQVPELV 707
Db 113 LYSQCRFDCCSLGVESAVGLAGVFTVTMASIAGLGLATWLGINFNAATTQIVPFLAL 172
QY 708 SV 709
Db 173 GI 174

RESULT 45
US-60-781-953-21
```

```
; Sequence 21, Application US/60781953
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jeffrey D.
; TITLE OF INVENTION: In Vivo Induced Genes of Mycobacterium tuberculosis
; FILE REFERENCE: 05-1122
; CURRENT APPLICATION NUMBER: US/60/781,953
; CURRENT FILING DATE: 2006-03-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 958
; TYPE: PRF
; ORGANISM: Mycobacterium tuberculosis
US-60-781-953-21

Query Match 1.6%; Score 111; DB 8; Length 958;
Best Local Similarity 19.7%; Pred. No. 7.5;
Matches 115; Conservative 60; Mismatches 195; Indels 214; Gaps 26;

QY 347 WVASWPLTILVSLVVPVVALAAGLVFTTDPVELMSAPNSQARSEKAFHQHGPFPFR 406
Db 28 WVA---LTIWNVVAP-----QLQSVARTHSVALGPHDA 58
QY 407 TNQVILTAPNRSSRYDSLLGPKNFSGILDLILLELELERLHLQVWSPRAQRNIS 466
Db 59 PSLIAMKRIGKDFQFDS-----DTTAMV-LLEGQEKL-----GDEAHR--- 96
QY 467 LQDICVAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANQTLMGQTSQVDMKDFLYCA 526
Db 97 FYDVLVTKLSQDTH-----VOHIEN-----FWGD----- 121
QY 527 NAPLTFKDGDTALALSCMADYGAPVFFFLAIGGYKGYSEARALIMTF-----SLNN 578
Db 122 ---PLT-----AAGSQSADGAAYVQLNLTGQGGSQANESVAAQRIVDSVPPPGIKA 173
QY 579 YPAGDPRLAQAKLWEEAFLEENRAFQRMAGMFQVTTAERLSLEDEINRTAEDLPAT 638
Db 174 YVTGPGPL-----GADRVVY-GDRSLH-----TITGISI--- 201
QY 639 SYIVFLVLSLALGSYSSWSRMVDSKATLGLGGVAVVLGVAMAGFFSVLGRS--- 694
Db 202 AVIAIMLFI-----AYRSLSAALI-MLLTVGL-----ELLAVRGIISTFAVNDLMGL 247
QY 695 SLVILQVVPFLVLSVGDNIFFIVLEYQRLPRRPGEPREVHICRALGRVAPSMMLCGLSE 754
Db 248 STFTVNVLTIAASTDYIIFLVGRYQE-ARATQGNREAAVYTFEGGTAHVVLASGLTV 306
QY 755 AICFFLGALTMPAVRTALTSGLAVIDFLQMSAFVALLSLDSKQRQASRLDVCCCVK 814
Db 307 AGAMYCLGFTRLPYFNTLASPCAIGLVTVMLASLTAPALIAVASR----- 352
QY 815 PQELPPQGGEGLLLGFFQKAVAPFLHWTGTVLL-----LFLALPGVSVLSYM 864
Db 353 -----FGLFDPKRAKTRRWRRTGTVVVRWPGVPLAATLLIALIGL----- 393
QY 865 CHISVGLQDELALPKDSYLLDYFLFNRYFEVGAPEVYFVTTLGY 908
Db 394 -----LALPK--YQTNV---NERY---IPSAAPSNI 419

RESULT 46
US-11-045-004-2027
; Sequence 2027, Application US/11045004
; GENERAL INFORMATION:
; APPLICANT: BUCHRIEGER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
```



```
Db 34 LGIAFOTLGVVYDGMGTSPLYVSDVFSKPIRSEVDVLGALSIVYITIAVIPLAKYVPV 93
Qy 706 VLSV-----GADNIFIFVLEYQRLPRPG-EPREVIHG-----RALG----- 741
Db 94 VLKANDGEGGTALYSILCRYAKVKNLFPNQPADQISSFRLKLPTELEALGIKEAL 153
Qy 742 --RVAPSMLLCSLSAICFFLICALTPMPAVRTFALTSGI-AVILDF--LLQMSAFVAL 794
Db 154 ETGKYLKTLILLVLMGTSMIIGDILTPAMSVMSAMSGLOGEVKGFGTNALVMSIVIL 213
Qy 795 LSLDSKRQASRLDVCCCKVQBELPPPGQEGLLGFFQKAYAPFLLHWTIRGVVLLPL 854
Db 214 VALFSIQRF-----GTGK---VGFL---FAPVLALW-----FF 240
Qy 855 ALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLPLNR-----YFEVGAPVYFVT----- 904
Db 241 SLGAIGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKQAMSGALGCVLCITGAEAM 294
Qy 905 --TLGYNFSSSEAGMNAICSSAGCNNSFTOKIQATEPPEQS-----YLAIPAS-SWVDDF 957
Db 295 FADLGHFSVRSIQMAFTCVVFPCLLLAYMGQAAYLTKEPEASARIFYDSVPKSLFWPVFV 354
Qy 958 IDWLTSSCCRLYISGPNKDKFCPSTVNSLNCNKMSTMGSVRPSVEQPHK----- 1010
Db 355 IATLAAMIASQAMISA-----TFSCVKQAMAL---GCFPRKLIHTSKRIGQ 399
Qy 1011 -YLP---WFLNDRPNIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEAL 1066
Db 400 IYIPVNWFL---MIMC-----ILV-----VSIF 420
Qy 1067 RAARELAANITADLRKVPGTDPAFEVFPYITITNVFYEQLTILPEGLFMLSCLVPTPAV 1126
Db 421 RSTTHIAN-----AYGIAEV-----GVMMVSTVLV----- 445
Qy 1127 SCLLGLDRLSGLNLLSIYMLIV-DTVGFMAL-WDISYNAVSLINLVSAGMSVE--FV 1182
Db 446 -----TLVMLLIQTNIFLALCFPLIFGSVETIYLLAVLTKILEGGWV 488
Qy 1183 SHITRSEALSTKPTW-----LERAKEA--TISM-----GSAVFAGVAMTNLPGILVLGL 1229
Db 489 PLVFATPFLTVMYIWNYSVLKYQSEVRERISMDPMBELGST-----LGTIRIPGI---GL 541
Qy 1230 AKAQILQ-----IFFRLNLLITLGLLHGLVFLPVLVSYVGPVNP 1271
Db 542 LYNELVQGISIFG---QFLLTLPALHSTIIF--VCIKYVPVPVVP 582

RESULT 50
US-11-056-355B-73132
; Sequence 73132, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 73132
; LENGTH: 827
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(827)
; OTHER INFORMATION: Ceres Seq. ID no. 4958916
US-11-056-355B-73132

Query Match 1.6%; Score 109.5; DB 7; Length 827;
Best Local Similarity 20.5%; Pred. No. 7.9;
```

```
Matches 145; Conservative 89; Mismatches 211; Indels 261; Gaps 41;
Qy 670 LGGVAVVLGAVMAAMG-----FFSYLGIRS-----SLVI--LQVVP-----FL 705
Db 89 LGIAFOTLGVVYDGMGTSPLYVSDVFSKPIRSEVDVLGALSIVYITIAVIPLAKYVPV 148
Qy 706 VLSV-----GADNIFIFVLEYQRLPRPG-EPREVIHG-----RALG----- 741
Db 149 VLKANDGEGGTALYSILCRYAKVKNLFPNQPADQISSFRLKLPTELEALGIKEAL 208
Qy 742 --RVAPSMLLCSLSAICFFLICALTPMPAVRTFALTSGI-AVILDF--LLQMSAFVAL 794
Db 209 ETGKYLKTLILLVLMGTSMIIGDILTPAMSVMSAMSGLOGEVKGFGTNALVMSIVIL 268
Qy 795 LSLDSKRQASRLDVCCCKVQBELPPPGQEGLLGFFQKAYAPFLLHWTIRGVVLLPL 854
Db 269 VALFSIQRF-----GTGK---VGFL---FAPVLALW-----FF 295
Qy 855 ALFGVSLYSMCHISVGLDOELALPKDSYLLDYFLPLNR-----YFEVGAPVYFVT----- 904
Db 296 SLGAIGIYNLLKYDFTVIRAL---NPFYIV---LFFNKNKQAMSGALGCVLCITGAEAM 349
Qy 905 --TLGYNFSSSEAGMNAICSSAGCNNSFTOKIQATEPPEQS-----YLAIPAS-SWVDDF 957
Db 350 FADLGHFSVRSIQMAFTCVVFPCLLLAYMGQAAYLTKEPEASARIFYDSVPKSLFWPVFV 409
Qy 958 IDWLTSSCCRLYISGPNKDKFCPSTVNSLNCNKMSTMGSVRPSVEQPHK----- 1010
Db 410 IATLAAMIASQAMISA-----TFSCVKQAMAL---GCFPRKLIHTSKRIGQ 454
Qy 1011 -YLP---WFLNDRPNIKCPKGLAAYSTSVNLTSDGOVLASRFMAYHKPLKNSQDYTEAL 1066
Db 455 IYIPVNWFL---MIMC-----ILV-----VSIF 475
Qy 1067 RAARELAANITADLRKVPGTDPAFEVFPYITITNVFYEQLTILPEGLFMLSCLVPTPAV 1126
Db 476 RSTTHIAN-----AYGIAEV-----GVMMVSTVLV----- 500
Qy 1127 SCLLGLDRLSGLNLLSIYMLIV-DTVGFMAL-WDISYNAVSLINLVSAGMSVE--FV 1182
Db 501 -----TLVMLLIQTNIFLALCFPLIFGSVETIYLLAVLTKILEGGWV 543
Qy 1183 SHITRSEALSTKPTW-----LERAKEA--TISM-----GSAVFAGVAMTNLPGILVLGL 1229
Db 544 PLVFATPFLTVMYIWNYSVLKYQSEVRERISMDPMBELGST-----LGTIRIPGI---GL 596
Qy 1230 AKAQILQ-----IFFRLNLLITLGLLHGLVFLPVLVSYVGPVNP 1271
Db 597 LYNELVQGISIFG---QFLLTLPALHSTIIF--VCIKYVPVPVVP 637

RESULT 51
US-60-752-355-33716
; Sequence 33716, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 33716
; LENGTH: 3979
; TYPE: prt
; ORGANISM: Streptomyces nanchangensis
US-60-752-355-33716

Query Match 1.6%; Score 109.5; DB 8; Length 3979;
Best Local Similarity 20.8%; Pred. No. 86;
Matches 99; Conservative 57; Mismatches 144; Indels 177; Gaps 25;
Qy 376 TTDPVELWASPNQARSEKAFHDQ-----HFGPFRTNQVI-----LTAPNRSS 419
```

Db 2850 TTEAERAMPPGAKAVDITGFDRAADGHHYGPSYGGQOTVMROGEDLLADITLPTACT 2909
QY 420 --YRYSLLGLGPNFSGIILDL--LLELLEQLERLHLQVMSPEAQRNISIQDI----C 471
Db 2910 PDHTTDSLAIHP-----ALLDAALHPLATADNPDG---EILWPTWSGVTLHATGATHV 2961
QY 472 YAPLNPNTSLYDCCINSLLQFQNNRTLLLL-----T 504
Db 2962 RARITPOGDNDYRLTLTD-----ATGQTVLTAGTIASRPDLTARLRTGPGDGLYQVRWT 3016
QY 505 ANQTLMGQTSQV--DK-----XDHFLYCANAPLTFKDGITALALSCNADYCAPVPEPL 554
Db 3017 AMPIPAGSATAVADDMWMLGDAGLKGGLADAVAPLASYPDVVAALVAAMD--GTPV--PSV 3074
QY 555 AIGGYKGYKDYSAEALIMTFSLNNYPAGDPRLAQAKL-----WEEAFLEE----- 599
Db 3075 VLTGLPADGGDADVVEVLTTAREWLAEPRLAESLVLVVTHDAVAEDTSDGPDGDDVD 3134
QY 600 -----MRAFO-----RRMAGM-----FQVTFARSL 621
Db 3135 PVAAGVWGLIRSAQSENPGRFTLLDLTRRDAGTAPDVVEVLRAAMDADWQVAVRGRAL 3194
QY 622 EDEINRTAEDLPIFATSVIIVIFLISALGYSYSSRSRVWDSKA----- 666
Db 3195 ---VPRLTAAD---AAAGIV-----LPVGA--PAWQLVMADERAGTVDGLAPECEPEVL 3240
QY 667 -TLGLGGVAVLGAVMAGMGPFSYLGIRSSLVLOVVP-----FLVLSVQAD 712
Db 3241 EPLAQGVQVIAVRA---AGVNF-----RDVMVTLGVDPDRRLGEGGAGWLDVAPD 3289

RESULT 52

US-60-752-355-39891
; Sequence 39891, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Adad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; PRIOR FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 39891
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Zea mays
US-60-752-355-39891

Query Match 1.6%; Score 108; DB 8; Length 474;
Best Local Similarity 19.2%; Pred. No. 4.5;
Matches 118; Conservative 76; Mismatches 210; Indels 212; Gaps 27;
QY 731 PREVH-----IGRALGRVAPSMMLCSLSAI-----CFFLGALTMPAVRTFALTGLAVI 781
Db 15 PRESHVAAEAKRLLSLAGPLVASCTIQNVQLVSMFVGHGLGELPLAGA-SLASSLANV 73
QY 782 LDPLQMSAFVALLSIDSKEASRLDVCCVKPQELPPGQEGEGL-----LLGPFQK--- 834
Db 74 TGFSLLVGM-----ASALDTLC-----GOAFGARQVGLGLYQKRAM 110
QY 835 -----AYAPFLHLWITRQVWLLFLALFGVSLYSNMCHISVGLDQSLALPKDSYLLDYFLF 889
Db 111 LVTLACVPIAAVWNAAGRILL-----LGQDRDIAEAGAYSRLILS 154
QY 890 LNRYPEVGAPVYFVT-----LGVNFSSEAGMAIC-----SSAGCNFSPFQKIQY 936
Db 155 LVPPVPLACHVRFLQTSIVPWPVMASSGATAGLHVLVCWALVFKAGMSGKAALSGAISY 214
QY 937 ATEPPEQSLAIAPASSWDDFDIMLTPSCCLLYISGPNKDF-----CPSTVNS 986
Db 215 SVNL-----AMLAL-----YVRLSSACKRTWTGFSFEAFRDLRLFTLAEVPSAM-- 258

QY 987 LNCLKNC---MSITGMSVRPSVEQPHKYLFWFLNDRPNI-KCPKGGLAAYSTSV-NLTS 1041
Db 259 MVCLEWSPFELIALLSGLLPNPKLETSVLISICLNTGALLFMVPHYGLCTAISTRVSNELGA 318
QY 1042 GOVLASRFWAYHKPLKNSQDYTEALRAARELANITADLRKVPDTPAPEVPPYTTNPV 1101
Db 319 GEPQAR-----LAARV----- 330
QY 1102 YEQYLITLPEGLFMLSCLVPTFAVSCLLGLDLRSGLNLNLSIVMLVDTVGFMALWDI 1161
Db 331 -----VNCIALSAGL-----LLGSTMILL-----RSFWGY 355
QY 1162 SY-NAVSLINLVSAVCMSEVFVSHITRSP--AISTKPTWLERAK-EATISMSGAVPAGVA 1217
Db 356 MYSNEPEVVTYIARM-MPVLAISPFDTGLHSLGSLGVLTCGRQKIGARVNLGAYILAGIP 414
QY 1218 MTNLPGLVLGLAKAQILQIPFPRNLNLTLLGLLHGLVFLDVLISYVGVDPVNPALALBQ 1277
Db 415 WA-----VLLAFVLHLNGMGLWGLVGVCSLTKLVLLMWTIRLN-----WE 455
QY 1278 KRAEBAVAAMVVASCP 1293
Db 456 KEATNAKETVPSSSLP 471

RESULT 53

US-10-724-972B-5410
; Sequence 5410, Application US/10724972B
; GENERAL INFORMATION:
; APPLICANT: DOUCETTE-STAMM, LYNN
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 47040.0007US
; CURRENT APPLICATION NUMBER: US/10/724,972B
; CURRENT FILING DATE: 2003-12-01
; PRIOR APPLICATION NUMBER: 09/134,001
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: 60/055,779
; PRIOR FILING DATE: 1997-08-14
; PRIOR APPLICATION NUMBER: 60/064,964
; PRIOR FILING DATE: 1997-11-08
; NUMBER OF SEQ ID NOS: 7546
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 5410
; LENGTH: 901
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-724-972B-5410

Query Match 1.5%; Score 107; DB 6; Length 901;
Best Local Similarity 17.4%; Pred. No. 14;
Matches 102; Conservative 104; Mismatches 239; Indels 142; Gaps 24;
QY 260 AASCPAIAARPOALDSTF-----YLGQMPG-SLVLIILCSVPVAVTILLVGFVRVAPAR 311
Db 207 AAGMPLIIISAIIGLSSVGIALLTYIPDNFTLLAVMIGLAVGIDYSLILFRFKEIK 266
QY 312 DKSRMVDPKGTSLSKLSFSFTHL-LQOPQGWGTWASWPLTILVLSVIPVALAAGL 370
Db 267 -----KKGVDVTVEAIIATAVGTAGSAVIFAGLTVMIACVGLSLGIDFLAVMGFASAI 318
QY 371 -----VPTLETDDPVELWSAPNSQARSEKAFHQHFGFPFRTNOVILLTAPNES-----SYR 421
Db 319 SVLFVAVLAALTLLPALI-----SIFHKSIIKKOKPTSKDKPKDHS 358
QY 422 YDSLILGPKNFSGIILDLLELLELLELLELLELLELLELLELLELLELLELLELLELLELLE 481
Db 359 WAKFIVGKPVIAVIVSL-IILILAAI-----PVSGKRLGIPDLSLKTDSSEYK 406
QY 482 LYDCCNISLLQYFQNNRTLLLL-----TANTLMGQTSQVDWKOHFLYCANAPLTF 532
Db 407 AYKLIISDNFGCY-NGQIVMLVNTKDGSGKSTIERDLANNRSDLEIDNVDTVSKAQLTD 465


```
; PRIOR APPLICATION NUMBER: 09/603,208
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
; PRIOR FILING DATE: 1999-07-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 306
; SEQ ID NO 272
; LENGTH: 783
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-10-703-799B-272

Query Match      1.5%; Score 105.5; DB 6; Length 783;
Best Local Similarity 20.0%; Pred. No. 15;
Matches 145; Conservative 108; Mismatches 268; Indels 205; Gaps 34;

QY 628 TTAEDLPATSVIVIFLYISLALSGSYSSWVMDSKATGLGVAHVILGVAAMAGFF 687
Db 194 TSEIIGIGIAFVILFTF-----GSLIAAGLPTAVIGVGIGALAIAT----- 240

QY 688 SYLGIRSSVILQVFPFLVSLVGADNIFVL-----EYQRLPRPGPREVHIGRALGR 742
Db 241 AFTDLNNVTPVLAVM--IGLAVGID-YALFILSRYAEYKMRP--ADAAGNAVGTAGSA 295

QY 743 V--APSMLLCSLSEAICFELGALTMPAVRTALTSGLAVIDFLQMSAFVALLSLDSK 800
Db 296 VVPAGATVIAIALVALIADIGFLTAM-----GISAAFTVFVAVIALIATFIPALGVFGG 349

QY 801 RQEAASLDVCCVKPQELPPGOG-----EGLL-----LGFFQKAYAPFLH 842
Db 350 HAFKGI-----PGLGNPTPKQTWEQALNRRSKGRSNWKLVRK--APGL-- 392

QY 843 WITRGVLLFLALFGVLSYMSCHISVGLDDELALPKDSYLLDYFLFLNRYFEVGAPVVF 902
Db 393 -----VVAWVVLGLGALTIPAM-----NLQLSLPSDS----- 419

QY 903 VTTLGNFSSEAGMNAICSSACGNFSTQKIQYATEPEQSYLAIPASSWDDPIDWLT 962
Db 420 TSNIDTQTSADLMAEGFGAGVN-----APPLVIVDTHEVNDSTALQ 463

QY 963 PSSCCRLYISGNKDKFCBPTSVNSLCLNCMSITMGSVRPSVEQFHKYLPWFLNDRNI 1022
Db 464 P-----LIEAQEPEEGEF-----DREQAARFATM--YVQTYS 496

QY 1023 KCPKGGLAAYSTSVNLTSDQVLSRPMAYHKPL-KNSQDYTEALRAARELAANITADLR 1081
Db 497 NIDVKNAQIISVNDDFTA-AQILVT---PYTGADKETPELMHVLRAQE-----AQIE 545

QY 1082 KYPGTD---PAFEVFPYITVVFYEQYLTILPEGLFMLSCLVPTFPFVASCLLG-----L 1133
Db 546 DVTGTELGTGTAVOLDIT-----SOLEDAPE-----VYLAWVGGLAIFLLVFRSLIV 596

QY 1134 DLRSGLMLLSI-----VMILVDTVGFMAIWD-----ISYNAVSLINLVSAVMSV--- 1179
Db 597 PLVAGLGLLSVGAAPGATVLVQSGFGFVNTPOPLISFMPFIPLIGTFGLANDYQVFL 656

QY 1180 -----EFVSHITRSPAISTKPTWLBRAKEATISMGs-----AVFAGVAMTNLPG 1223
```

```
Db 657 VTRMREHYTHNGKQPGSKYTPVEQSVIEGTQGSRVVTAALIMIAVFVAFIDQPLPF 716
QY 1224 ILVLG--LAKAQLIQIPFFRLNMLITLLGLL-HGLVFLPVLVSYPGPDVN-PALAEOKR 1279
Db 717 IKIFGALGAGVFFDAFFIRMLGVPASMFMLGKATWMPKWLDRILPDSIDIEGTALEKEW 776
QY 1280 ABEAVA 1285
Db 777 BEKQAA 782

RESULT 62
US-60-751-420-2802
; Sequence 2802, Application US/60751420
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Van Berdewegh, Paul
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: GeneMap of the human genes associated with Crohn's disease
; FILE REFERENCE: GENI-011/00US
; CURRENT APPLICATION NUMBER: US/60751.420
; CURRENT FILING DATE: 2005-12-19
; NUMBER OF SEQ ID NOS: 27266
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2802
; LENGTH: 1077
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-60-751-420-2802

Query Match      1.5%; Score 105.5; DB 8; Length 1077;
Best Local Similarity 18.9%; Pred. No. 25;
Matches 167; Conservative 93; Mismatches 277; Indels 347; Gaps 40;

QY 607 MAGMFQVTFTAERSLDEINRTAEDLPATSVIVIFLYISLALSGYS-----SMSR- 659
Db 41 LAALLAVAWASGRELTSD-----PSFLTTLVLCALGGFSLLLGLASRQRQRWTRP 91

QY 660 ----VMYDSKATLGLGVAVVILGAVMAAGFFSYLGRSSVILQVFPFLVSLVGADNIF 715
Db 92 LSGLVWV---ALLALGHAFILFTGGVSAWDOVSF-----LF 125

QY 716 IFVLEYORLPRRPGPREVHIGRALGR--VAPSMMLCSLSEAIC--PFLG-----ALT 764
Db 126 VIFTAYAWLP-----LGMDDAAVAGLASSLSHLLVILGVLGPQPSRPALL 171

QY 765 PMPAVRTALTSGLA-----VILDFLQMSAFVALLSLDSKR-----QEAASLDVC 810
Db 172 POLAANAVFLFCGNVAGVYHKALMERALRATFREALSSLSHRRRLDTEKKHQEHLLSIL 231

QY 811 CCVKPQELP-----PPGOG-----EGU-----LLGFFQKAYAPFL 840
Db 232 PAYLAREKAEIMARLQAGQGSRPSTNNFSLHYVKHQGVSVLYADIIVGTR-----L 285

QY 841 LHMTRGVLLFLALFG-----VSLYSMCH-----ISVG 870
Db 286 ASECSPKELVLMNELFGKFDQIAKEHECWRIKILGDCYCVSGCLPLSLPDHAINCVRMG 345

QY 871 LDQ-----ELALPKDSY-LLDYFLFLNRYFEVGAPV 900
Db 346 LDMCRAIRKLRAATGVDDINMRVGHSGSVLCGVIGLQKQWQVMSHSDVTLANHMEAGVVP 405
```

QY 901 YFTTILGVNFSBAGNAICSSAGCNPFSTQKIYATFPFQSYLAIPASSWDDPDIW 960
Db 406 GRVHTIGATALLAGAYAV-EDAGMEH-----RDPYLRELGPFTYVIDPRAEED-----455
QY 961 LTPSSCCRLYISGPNKDKFCPSTVNSLNCIKCMISITM-----GSVRPSVEQFHKYL 1013
Db 456 -----EKGTAGLLSLEGLKMRPSLLMTRYLESWGAAPPAHLSH-----496
QY 1014 WFLNDRP---NTKCPKGGIAAYSTSVNL-----TSDG---QVLASRFMAYH 1053
Db 497 ---GDSVSTSTPLPEKTLASFTQWSLDRSRTPRGLDDDELDTGDAKPPQVIEQ---LNSQ 551
QY 1054 KPLKNSQDYEARALARELAANITADLRKVPGTDAPE-----VPPYTTITNVPEOYLTI 1108
Db 552 KQWKOSKOPNPLTYFREKEMEKEVRLSAI-----PAFKYEAECTFLVFLSNFIQMLVTN 607
QY 1109 LPEGL-----FMLSCLIV-----PTFAVSCLLG-----LDRSGLLN-----1141
Db 608 RPPALAITYSITFLLELILFCFSEDLMRCVLKGPMLHWLPALSGLVATRPGLRIALG 667
QY 1142 ----LLSIVMLVDTVGFMAWDISYNAVSLINLSAVGMSVEFVSHITRSPAISTKPTW 1197
Db 668 TATILLVPAMATISLPPFPPTSDCPQAPNVSSMISNL-----VLGLAKAQIQLIIPFRLNLL 1245
QY 1198 LERAKEATISGMSAVFAGVAMTNLPGIL-----ELGSLPLISVPYSMHCTILGFLSCSLFLHMSFELKIL 745
Db 708 -----LHGLVFLP---VILSVVGP-DYNPALALROK 1278
QY 1246 LLLLMLAASCSLFLHSHAWLSECLIVRLYLPLDRGVLKEPK 789
Db 746 LLLLMLAASCSLFLHSHAWLSECLIVRLYLPLDRGVLKEPK 789
RESULT 63
US-11-360-355-119639
; Sequence 119639, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijing
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: 38-21(53885)
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 119639
; LENGTH: 308
; TYPE: prt
; ORGANISM: Heterodera glycines
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (121)..(121)
; OTHER INFORMATION: Xaa can be any naturally occurring amino acid
; FEATURE:
; OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 46062; Strand=-; Position=1
; OTHER INFORMATION: -140,252-307,384-569,654-752,801-998,1285-1533
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit_ID=NP_494986.1; Match level="QueryCoverage=100%, HitCoverage=29%, E-value=5e-81, Identity=50%"; Hit description: Patched Related (ptr-6) [Caenorhabditis elegans] gb|AAA68302.1
; OTHER INFORMATION: Hypothetical protein C54A12.1 [Caenorhab
; FEATURE:
; OTHER INFORMATION: Essential gene: C elegans homolog=C54A12.1; Phenotype=Emb
US-11-360-355-119639

Query Match 1.5%; Score 105; DB 7; Length 308;
Best Local Similarity 18.9%; Pred. No. 4.1;
Matches 39; Conservative 29; Mismatches 66; Indels 72; Gaps 6;
QY 746 SMLCSSLSEACFPGLGALTPMPAVTEPALTSLGLAVILDFELLQMSAFVALLSLDSK--RQE 803
Db 5 AIPFTISITDVSFAAGTYSIDILAVRGPCMTAACMMFTLYQITTPAALMWISSRLQMOQ 64
QY 804 ASRLDVC--CC-----
Db 65 RNACNACLPCCLLAKDPPEFEBEGKTELNNNNNNKNDKSSKAMDYEERDEKGFQKKIKN 124
QY 816 QELPPPGQEGULL-----GPFQKAYAPFLLHWTTRGVVLLFLA 855
Db 125 NTSAPISISSSSSALSASEFNLAMRQKPGMKFRNFYVNFYVDFLDMRTKFAVGTIPFV 184
QY 856 LFGVSLYSNMCHISVGLDOR-LALPKD 880
Db 185 YLAISLWGIIVSMQ-GLDYEKLLSSD 209
RESULT 64
US-11-056-355B-69790
; Sequence 69790, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nickolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69790
; LENGTH: 1080
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1080)
; OTHER INFORMATION: Ceres Seq. ID no. 1825823
US-11-056-355B-69790
Query Match 1.5%; Score 105; DB 7; Length 1080;
Best Local Similarity 19.7%; Pred. No. 27;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;
QY 300 ILLVGRVA-PARDKSKMVDPKKGTSLSDKLSFSTHTLLGQPFQGMG-----TWVASWP 352
Db 44 VOLIGFPVLFLFRPFSQTKNPK--PTEADFRKFSSTILGSVYIVTGLLVANSYSSVG 101
QY 353 LTLVLVSVVVALAAGLVFTTELTDTPVELWSAPNSQARSEKAFHDQHPGPPFRTNQVIL 412
Db 102 LLYLFVSTFSLI-LASQLAFT-----APFSYFLNSQKFTPPVNSLFL 144
QY 413 TAPN-----RSSRYDLSLLGPKNPSGIL-----DLDLLELLEL-QERLRHLQVWSP 459
Db 145 TISSALLVNTDSENTAKVRVKYVIGICTIGASAGIGLLSLVQLILRKVKKQTPS- 203
QY 460 EAQRNISLQD-ICYAPLNPDTSLYDCCINSLLQYFQNN--RTLLLTANOTLMGQTSQV 516
Db 204 -----TWTDLVAY-----QSLVASCV-VLIGLFASGEWKTLSSEMYKL-GKVPV 248
QY 517 -----DWKDFHYCANAPLTFKDGITALALSCHADYGAVPFPLAIGGYKKG-DYSEA 567
Db 249 MTLASIAISQWYTTIGVVG--LIFESSSVFSNSITA-VGLPTVPVAVIVPHDKWASKI 305
QY 568 EALIM-----TFSLNYY-----PAGDPRLAQAKLWEEAFLEEMFAQRBMAG 609
Db 306 FSIILAIWGFISVYQHYDLDEKLLKTSHTSPVGDPHLLPA-----EEG-----348

```
QY 610 MFQVTTAERSLE-----DEINRTTAEDLPATSYIVIFLYISLALGSSYSWRVMVD 663
Db : : : : :
349 --HTNIHQDLEANLIDHEVVTSSSAVP-----QTEYKRWLRVSIY 391
QY 664 SKATLGLGGVAVVLGAVMAAGFFS--YLGIRSSLVILQVVPFLVLSVGADNIPFIVLEQ 722
Db : : : : :
392 VIFVLFCQPLATILGRLYYENGSTYV-----VTLLQLIGFPLV-----LFRF---FS 438
QY 723 RLPRRGEPREVHIGRALGRVAPSMLLCSLSEBAICFLGALTMPAVRTFALTS-----GL 778
Db : : : : :
439 RIQPKSTDTNFSQSPSFTTLASVYLCTGLLSAVAYLSALA--FTAFPSYFLNSQKFTPL 497
QY 779 AVILDFLLQMSARVALLSDSKQ--EASR-----LDVCCCKVQBELPPQGGGLLLGFPQ 833
Db : : : : :
498 IVSSLLLLTVSSALLVNTDSENSTNVRQVYVIGFICTI-----GASAGIGLLLSLIQ 551
QY 834 KAYAPFLHWITRGVVLLEFLALFGVSLYSMCHISVGL-----DOELALPKDS 881
Db : : : : :
552 MLPRKVTKHTSSAVT---DLAIY-QSLVASCVVILGLFASGEWETLPSEMNYKLGKVS 607
QY 882 YLLD-----YFL-----FLNRYFEVGAPVY----- 901
Db : : : : :
608 YVLTLASAAISQVYTLGLVGLIFESSVSFSNSITAVGLPIVPAVAIVFHDMDASKIF 667
QY 902 -----FVTLTGVNFSSEAGMNAICSSA-----GCNFSFTQKIQ--YATERPE 942
Db : : : : :
668 SIILATCGFLSPYQHYLDEKLTNTSHTSAVGDLHLFVEEGHTNIQSDQLEANLLDHEE 727
QY 943 QSYLAIPASS---WV-----DDFIDMLTPSSCC--RLYISGPNKDKFCPSVNSLNCIK 991
Db : : : : :
728 TESFSVPQTKNCKRWLRVSIYAIFFVFCQPLATVLRGLYENGKSTYVVTLLQLIGFPV 787
QY 992 NCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db : : : : :
788 LILFRFESRIQPKSTDTNFSQSPSFTTLASVYLCT--GLLSAVAYLSAVGLLPLVST 845
QY 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAFVFPYIT 1098
Db : : : : :
846 FSLILASQLAFTAFPSYFLNSQKFTPLIVNSILFLATVSSALL--VVNTDSE-----NTY 897
QY 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC--LL 1130
Db : : : : :
898 NVSRVQVYVIGFICTIGASAGIGLVLSLIQLLPRKVTKHTSSAVLDLANYQSLVATCVVL 957
QY 1131 LGL-----DLRSGLLMLSIVMLVDTVGFMALWDI--SYNAVSLI-----N 1170
Db : : : : :
958 IGLFASGEWRTLPSEMRNYKLGKVSIIYLTLASAAIP---WQVYTVGCVLIFESSVSFSN 1014
QY 1171 LVSAVGMV 1179
Db : : : : :
1015 SITAVGLPI 1023
```

RESULT 65

```
US-11-056-355B-69789
; Sequence 69789, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69789
; LENGTH: 1097
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
```

```
; LOCATION: (1)..(1097)
; OTHER INFORMATION: Ceres Seq. ID no. 1825822
US-11-056-355B-69789
```

```
Query Match 1.5%; Score 105; DB 7; Length 1097;
Best Local Similarity 19.7%; Pred. No. 28;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;
```

```
QY 300 ILLVGRVA--PAROKSMVDPKGTSLSDKLSFSHTLLGQFFQGWG-----TWVASWP 352
Db : : : : :
61 VOLIGFVLFLRPFSSQTKNPK--PTADPRKSSFTILGSGVIYVTGLLSANSVMSVG 118
QY 353 LTILVLSVIPVVAALAGLVFTTELTPDVELMSAPNSQARSEKAPHQDHGFFFTNQVIL 412
Db : : : : :
119 LLYLPVSTFSLI--LASQLAFT-----AFFSYFLNSQKFTPIVNSLFL 161
QY 413 TAPN-----RSSRYDLSLLGPKNFSGIL-----DLDLLELLEL--QERLRHLQVSWP 459
Db : : : : :
162 TISSALLVNTDSENTAKVSRVKVIGICTIGASAGIGLLLSVLQLILRKVLKKQTFP- 220
QY 460 EAQRNLSIQD--ICYAPLNPDNTSLYDCINSLLQYFQNN--RTLLLTANQTLMGQTSQV 516
Db : : : : :
221 -----TVTDLVAY-----QSLVASCV-VLIGLFASGEWKTLTSEMENYKL--GKPYV 265
QY 517 -----DWKDFLYCANAPLTFKDTALALSCHADYGAFFPLAIGGYKKG--DYSEA 567
Db : : : : :
266 MTLASIAISMQVYTVGVG--LIFESSVSFSNSITA--VGLPIVPAVAIVFHDMDASKI 322
QY 568 EALIM-----TFSLNNY-----PAGPRLAQAOKLWEEAFLEEMRAFQRRMAG 609
Db : : : : :
323 PSIILATWGFISFYQHYLDEKLTNTSHTSPVGDPHLLPA---EEG----- 365
QY 610 MFQVTTAERSLE-----DEINRTTAEDLPATSYIVIFLYISLALGSSYSWRVMVD 663
Db : : : : :
366 --HTNIHQDLEANLIDHEVVTSSSAVP-----QTEYKRWLRVSIY 408
QY 664 SKATLGLGGVAVVLGAVMAAGFFS--YLGIRSSLVILQVVPFLVLSVGADNIPFIVLEQ 722
Db : : : : :
409 VIFVLFCQPLATILGRLYYENGSTYV-----VTLLQLIGFPLV-----LFRF---FS 455
QY 723 RLPRRGEPREVHIGRALGRVAPSMLLCSLSEBAICFLGALTMPAVRTFALTS-----GL 778
Db : : : : :
456 RIQPKSTDTNFSQSPSFTTLASVYLCTGLLSAVAYLSALA--FTAFPSYFLNSQKFTPL 514
QY 779 AVILDFLLQMSARVALLSDSKQ--EASR-----LDVCCCKVQBELPPQGGGLLLGFPQ 833
Db : : : : :
515 IVSSLLLLTVSSALLVNTDSENSTNVRQVYVIGFICTI-----GASAGIGLLLSLIQ 568
QY 834 KAYAPFLHWITRGVVLLEFLALFGVSLYSMCHISVGL-----DOELALPKDS 881
Db : : : : :
569 MLPRKVTKHTSSAVT---DLAIY-QSLVASCVVILGLFASGEWETLPSEMRNYKLGKVS 624
QY 882 YLLD-----YFL-----FLNRYFEVGAPVY----- 901
Db : : : : :
625 YVLTLASAAISQVYTLGLVGLIFESSVSFSNSITAVGLPIVPAVAIVFHDMDASKIF 684
QY 902 -----FVTLTGVNFSSEAGMNAICSSA-----GCNFSFTQKIQ--YATERPE 942
Db : : : : :
685 SIILATCGFLSPYQHYLDEKLTNTSHTSAVGDLHLFVEEGHTNIQSDQLEANLLDHEE 744
QY 943 QSYLAIPASS---WV-----DDFIDMLTPSSCC--RLYISGPNKDKFCPSVNSLNCIK 991
Db : : : : :
745 TESFSVPQTKNCKRWLRVSIYAIFFVFCQPLATVLRGLYENGKSTYVVTLLQLIGFPV 804
QY 992 NCMSITMGSVR--PSVEQFHKYLPMFLNDRPNIKCPKGGLA-----AYSTSVNL-----TS 1040
Db : : : : :
805 LILFRFESRIQPKSTDTNFSQSPSFTTLASVYLCT--GLLSAVAYLSAVGLLPLVST 862
QY 1041 DGOVLASR--FMAYHKPLKNSQDYTEALRAARELAANIATDLRKVPGTDPAFVFPYIT 1098
Db : : : : :
863 FSLILASQLAFTAFPSYFLNSQKFTPLIVNSLFLATVSSALL--VVNTDSE-----NTT 914
QY 1099 NVFYEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC--LL 1130
```

Db 915 NVSRVQVIGFCTTGASAGIGLVLSLQLLPKVKFTKHTSSAVLDLDANYQSLVATCVWL 974
Qy 1131 LGL-----DLRGLNLLSIVMLVDTVGFMALWDI-SYNVSLI-----N 1170
Db 975 IGLPASGEWRTLPSEMRNYKLGKVSYLTLASAAIF--WQVYTVGCVLFPSSSVFSN 1031
Qy 1171 LVSAVMSV 1179
Db 1032 SITAVGLPI 1040

RESULT 66
US-11-056-355B-69788
; Sequence 69788, Application US/11056355B
; GENERAL INFORMATION:
; APPLICANT: Brover, Vyacheslav
; APPLICANT: Alexandrov, Nikolai
; TITLE OF INVENTION: Sequence Determined DNA Fragments and Corresponding
; FILE REFERENCE: 2750-1590PUS2
; CURRENT APPLICATION NUMBER: US/11/056,355B
; CURRENT FILING DATE: 2005-02-14
; PRIOR APPLICATION NUMBER: 60/544,190
; PRIOR FILING DATE: 2004-02-13
; NUMBER OF SEQ ID NOS: 119966
; SEQ ID NO 69788
; LENGTH: 1128
; TYPE: prt
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: peptide
; LOCATION: (1)..(1128)
; OTHER INFORMATION: Ceres Seq. ID no. 1825821
US-11-056-355B-69788

Query Match 1.5%; Score 105; DB 7; Length 1128;
Beat Local Similarity 19.7%; Pred. No. 29;
Matches 215; Conservative 159; Mismatches 397; Indels 318; Gaps 59;

Qy 300 ILLVGFRA-PARDKSKWDPKKGYSLSGKLSFSTHTLLGQFFQWG-----TWVASWP 352
Db 92 VOLIGFPLFLFRFFSQTKNPK--PTEADFRKFSFTILGVSIVITGLLVANSYMSV 149
Qy 353 LTLVLSVIVVVALAAGLVFTLTDPVLSAPNSQARSEKAFHDQHFQPPFRNQVIL 412
Db 150 LYLVPVSFSLI-LASQAF-----AFFSYFNSQKFTPIVNSLFL 192
Qy 413 TAPN-----RSSRYDLSLLGPNFSGIL-----DLDLLELEL-OERLRHLQVWSP 459
Db 193 TISSALLVVNTDSENTAKVSRVKYVIGIITIGASAGIGLLLSVLQILRKVKKQTF- 251
Qy 460 EAQRNLSQD-ICYAPLPNDTSLYDCINSLLQFQNN--RTLLLTANQTLMOQTSQV 516
Db 252 -----TVDLVAY-----QSLVASCV-VLIGLPASGEWKTLTSEMYNKL-GKVPYV 296
Qy 517 -----DWKDHLYCANAPLTKDGTALALSCMADYCAPVFPFLAIGGYKGK-DYGEA 567
Db 297 MTLASIAISQWYTTGVVG--LIFESSVSFNSITA-VGLPIPVVAVIVFDKWNASKI 353
Qy 568 EALIM-----TFSLNNY-----PAGDPRLAQAKLWEBAFLMEARAFORMAG 609
Db 354 FSIILAIWGFISFVQHYLDKELKTSHTSPVGDPHLLPA---EEG----- 396
Qy 610 MEQVFTAESLE-----DEINRTAEDLPFPATSYIVIFLYISLALGSYSSWSRVMD 663
Db 397 ---HTNIHSDQNLEANLIDHEVVTSSSAVP-----QTENYKRWLRVSIY 439
Qy 664 SKATIGLGGVAVLGVAMAAAGFFS-YLGIIRSSLVQLQVVPFLVSVGADNIFIFVLEQ 722
Db 440 VIFVLCQPLATILGRLYYENGNTYV-----VTLLQLIGFPVLV-----LFRF---FS 486
Qy 723 RLPRPGPREVHIGRALGRVAPSMLLCSLSAICFFLGLTTPMPAVRTFALTS-----GL 778

Db 487 RIROPKSTDTNFSOSPSTFTLASVYVLTGLLVSAYAYSALA-FTAFSPYPLNSOKFTPL 545
Qy 779 AVILDFLQMSAFVALLSLDSKRO-EASR-----LDVCCCVKQQLPPLPQOQGLLGGFP 833
Db 546 IVSSLLLTIVSALLVNTDSENSNVSRVQVVGICTI-----GASAGIGLLSLIQ 599
Qy 834 KAYAPFLHWITRGVVLFLFALFGVSLYSMCHISVGL-----DQELALPKDS 881
Db 600 MLFRKVFTKHTSSAVT---DLAIY-QSLVASCVLIGLPASGEWETLPSEMRNYKLGKVS 655
Qy 882 YLLD-----YFL-----FLNRYFVCGAPVY----- 901
Db 656 YVLTLASAAISQWYVTLGLVGLFPSSSVFSNSITAVGLPIVPAVAIVFHDRMDASKIF 715
Qy 902 -----FVTTLGYNFSSEAGMNAICSSA-----GCNFSFTQKIQ-YATEPPE 942
Db 716 SIILAIACGLFVQHYLDKELKTSHTSASVGDHLHPVEEGHTNTQSDNLEANLIDHEE 775
Qy 943 QSYLAIPASS---WV-----DDFIDMLTPSSCC--RLYISGPNKDKFCPSTVNSLNCILK 991
Db 776 TESFSPQTKNCKRWLRVSIVAIFVIFCQPLATVLGRLYYENGKSTVYVTLQLIGPPV 835
Qy 992 NCMSTMGSVR--PSVEQFHKYLPWFINDRPNIKCPKGLA-----AYSTSVM- 1040
Db 836 LILFRFPRIRQPKSTDTNFSQSPSTFTLASVYVLT--GLLVSAAYAYSALVGLLPLVST 893
Qy 1041 DQOVLASR--FMAYHKPLKNSQDYTEALRAARELANITADLRKVPDPAPEVFPYTT 1098
Db 894 FSLILASQLATFAPFSYFLNSQKFTPLIVNSLFLTSSALL--VVNTDSE-----NTT 945
Qy 1099 NVFVEQYL-----TI-----LPEGLF-----MLSLCLVPTFAVSC-LL 1130
Db 946 NVSRVQVIGFCTTGASAGIGLVLSLQLLPKVKFTKHTSSAVLDLDANYQSLVATCVL 1005
Qy 1131 LGL-----DLRGLNLLSIVMLVDTVGFMALWDI-SYNVSLI-----N 1170
Db 1006 IGLPASGEWRTLPSEMRNYKLGKVSYLTLASAAIF--WQVYTVGCVLFPSSSVFSN 1062
Qy 1171 LVSAVMSV 1179
Db 1063 SITAVGLPI 1071

RESULT 67
US-11-360-355-129465
; Sequence 129465, Application US/11360355
; GENERAL INFORMATION:
; APPLICANT: Boukharov, Andrey
; APPLICANT: Du, Zijiang
; APPLICANT: Guo, Liang
; APPLICANT: Kovalic, David
; APPLICANT: Lu, Maolong
; APPLICANT: McCarter, James
; APPLICANT: Miller, Nancy
; APPLICANT: Williams, Deryck
; APPLICANT: Vaudin, Mark
; APPLICANT: Wu, Wei
; TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS
; FILE REFERENCE: IN PLANTS AND COMPOSITIONS THEREOF
; CURRENT APPLICATION NUMBER: US/11/360,355
; CURRENT FILING DATE: 2006-02-24
; NUMBER OF SEQ ID NOS: 171306
; SEQ ID NO 129465
; LENGTH: 122
; TYPE: PRP
; ORGANISM: Heterodera glycines
; FEATURE:
; OTHER INFORMATION: Coding regions on cDNA: vcdna=SeqID_55888; Strand=-; Position=1
; OTHER INFORMATION: -91,165-245,660-825
; FEATURE:
; OTHER INFORMATION: Homolog annotation: Hit_ID=NP_491005.2; Match level="QueryCovera

OTHER INFORMATION: =95%, HitCoverage=9%, E-value=7e-31, Identity=60%, Hit description
US-11-360-355-129465

Query Match 1.5%; Score 104; DB 7; Length 122;
Best Local Similarity 24.1%; Pred. No. 1.2;
Matches 27; Conservative 27; Mismatches 38; Indels 20; Gaps 3;

QY 698 ILQVPELVLSGADNIFVLEYQRLP-----RRP-----GEPREHIG----- 737

DB 10 VLVCTPFLVLAIGVDVDAIMHAWQRITDQWRKPHAWQRITDQWRKPIGNISCSKLP 69

QY 738 RALGRVAPSMILCSSEALICFPLGALTMPAVRTFALTSLGLAVILDFLLQMS 789

DB 70 KVLVETGPAVLISALTNICADAVGSGFTGSPETILCVGNMAIAVDPIYQIT 121

RESULT 68

US-11-360-355-149551
Sequence 149551, Application US/11360355

GENERAL INFORMATION:

APPLICANT: Boukharov, Andrey

APPLICANT: Du, Zifeng

APPLICANT: Guo, Liang

APPLICANT: Kovalic, David

APPLICANT: Lu, Maolong

APPLICANT: McGarner, James

APPLICANT: Miller, Nancy

APPLICANT: Williams, Deryck

APPLICANT: Vaudin, Mark

APPLICANT: Wu, Wei

TITLE OF INVENTION: METHODS FOR GENETIC CONTROL OF HETERODERA INFESTATIONS

TITLE OF INVENTION: IN PLANTS AND COMPOSITIONS THEREOF

FILE REFERENCE: 38-21(53885)

CURRENT APPLICATION NUMBER: US/11/360,355

CURRENT FILING DATE: 2006-02-24

NUMBER OF SEQ ID NOS: 171306

SEQ ID NO 149551

LENGTH: 238

TYPE: PRT

ORGANISM: Heterodera glycines

FEATURE:

NAME/KEY: misc feature

LOCATION: (186)..(186)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

NAME/KEY: misc feature

LOCATION: (194)..(194)

OTHER INFORMATION: Xaa can be any naturally occurring amino acid

FEATURE:

OTHER INFORMATION: Coding regions on vcdna: vcdna=SeqID 75974; Strand=-; Position=1

OTHER INFORMATION: -114,485-520,583-627,956-1094,1610-1739,2041-2109,2471-2644

OTHER INFORMATION: Homolog annotation: Hit ID=CAE60539.1; Match level="QueryCoverage

OTHER INFORMATION: =9%, HitCoverage=25%, E-value=2e-54, Identity=50%; Hit description

OTHER INFORMATION: =Hypothetical protein CBG04166 [Caenorhabditis briggsae]

US-11-360-355-149551

Query Match 1.5%; Score 104; DB 7; Length 238;

Best Local Similarity 21.6%; Pred. No. 3.3;

Matches 47; Conservative 44; Mismatches 87; Indels 40; Gaps 10;

QY 556 IGGYKGKDYSEALIMTFSLNNYPAGDPLRAQAKLWEEAFLEENRQRMRMAGHFWQTF 615

DB 27 IGGWKDDISDYE-----MSISNY-----FKDKYTGQ-----KYIGQHIVL 63

QY 616 TAERSLEDEINRTAEDLPFATSY-----IVIFLVISALGSGYSWSKVMVD-SKATLGL 670

DB 64 TISTSVQVEVDAGN-----IIRSFGLGLAWMLCSLFSNLSAFFPHQFSYIKLPVAI 118

QY 671 GGVAVVLGVMAAMGFFSYLGRSSILVQLQVVPFLVLSVGADNIFIVLEYQ---RLPRR 727

DB 119 FACLCPPFNASGTALGILFPAGVRHS-SILAVTFPILALGVDDAFIMTHSWQLASKRRRH 177

QY 728 PGEPEVH--IGRALGR-VAPSMILCSLSEAICFPLGA 762

DB 178 EGVANDANKLINQSDXNDGTGPAILISTUTNIFADLVGA 215

RESULT 69

US-10-724-972B-4902

Sequence 4902, Application US/10724972B

GENERAL INFORMATION:

APPLICANT: DOUCETTE-STAMM, LYNN

APPLICANT: BUSH, DAVID

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 47040.0007US

CURRENT APPLICATION NUMBER: US/10/724,972B

CURRENT FILING DATE: 2003-12-01

PRIOR APPLICATION NUMBER: 09/134,001

PRIOR FILING DATE: 1998-08-13

PRIOR APPLICATION NUMBER: 60/055,779

PRIOR FILING DATE: 1997-08-14

PRIOR APPLICATION NUMBER: 60/064,964

PRIOR FILING DATE: 1997-11-08

NUMBER OF SEQ ID NOS: 7546

SOFTWARE: PatentIn version 3.3

SEQ ID NO 4902

LENGTH: 510

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

US-10-724-972B-4902

Query Match 1.5%; Score 104; DB 6; Length 510;

Best Local Similarity 22.2%; Pred. No. 11;

Matches 50; Conservative 42; Mismatches 93; Indels 40; Gaps 7;

QY 985 NSLCLNKMCSITMGSVRPSVEQFKYLPWFNLDRPNIKCPKGGLAAYSTSVNL----- 1038

DB 308 NSLSEPNADSYGSG-----WFINTDHLVFTGTVDNFSQILLNIRKSY 353

QY 1039 -----TSDQVLAERFAYH--KPLKNSQDYTEALRAARELANITADLRKVGTPDA 1089

DB 354 GIVVLANTNSQVTR---LAEHLNTQIMNRRHYTTI-----EEKVNOTKDMQLIIST--L 403

QY 1090 FEVPYITITNVFYEQYLTILPEGLFMLSCLVPTFAVSCLLGLDLRGLNLLSLVMIL 1149

DB 404 ADIFWVIFSIILVFSKILK-LREGHIFIRKCLRTSIMPSSIILGTFVAMMILFYLLFLIIG 462

QY 1150 VDTVGFMAWDISYNAVSLINLVSAGVMSVEFVSHITRSPAIATK 1194

DB 463 DATWGFVLSWLPKLSKYLVSVLAITMLLVLSLISITYRSDKK 507

RESULT 70

US-10-159-257C-191

Sequence 191, Application US/10159257C

GENERAL INFORMATION:

APPLICANT: SHEN, BEN

APPLICANT: LIU, WEN

APPLICANT: CHRISTENSEN, STEVEN D.

APPLICANT: STANDAGE, SCOTT

TITLE OF INVENTION: GENE CLUSTER FOR PRODUCTION OF THE ENEDIYNE ANTITUMOR

FILE REFERENCE: 407T-896020US

CURRENT APPLICATION NUMBER: US/10/159,257C

CURRENT FILING DATE: 2002-05-31

PRIOR APPLICATION NUMBER: 09/478,188

PRIOR FILING DATE: 2000-01-05

PRIOR APPLICATION NUMBER: 60/115,434

PRIOR FILING DATE: 1999-01-06

NUMBER OF SEQ ID NOS: 225

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 191

LENGTH: 692

Db 60 AKNPVAGPKDGKSIIFRELALFQDYHGLPSFKKALVDFMABIRGNKVTFDPNHIIVLTAG 119
QY 416 NRSSVRY-----DSELLGPKNPSGILDLDL-----LELELQ-----ERLRH 453
Db 120 STSANQTLMFCLAEIGDAFLPTPYPG-FDRDLKRTGVEIIVPTQCTSSNNPQTAEAL 178
QY 454 LQWSPQAQRNLSLODICYA-PLNPNTSLYDCCINSLQYQNNRTLLALLTANQTLMOQ 512
Db 179 KQAYQBAKRNLRVGLVTNPNPLGTTMSRSELNLLVDFIKKN--MHLISDEIYSOT 236
QY 513 T-----SOVDWKDHFLYC-ANAPLTFKDGDTALALSCWADYGAIPFPFLAIG 557
Db 237 VYNSPGFVSIMBILKDRNLXD-----CGANV-----WDRVHIYVLSKDLGLPGFRVGAI- 287
QY 558 GYKGDYSEAEALI-----MTFSLNYPAGDPRLAQAOKLWEEAFL-----BEMRA 602
Db 288 -----YSENEVVAAATKSSFGVLVSQTY-----LLSAMLGDKFKTKYIYSENLR 335
QY 603 FORM-----AGMF-----QVTFAPERSLEDEI 625
Db 336 LKRRQRNLVSGLOKAGISCLKTNNAGLFCWDMRHLHLSNTPEAEMLWKI 387

RESULT 75

PCT-US06-00964-7153
; Sequence 7153, Application PC/TUS0600964
; GENERAL INFORMATION:

; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: PCT/US06/00964
; CURRENT FILING DATE: 2006-01-23
; NUMBER OF SEQ ID NOS: 19250

; SEQ ID NO 7153

; LENGTH: 618

; TYPE: PRT

; ORGANISM: Oryza sativa (japonica cultivar-group)

PCT-US06-00964-7153

Query Match 1.5%; Score 103.5; DB 1; Length 618;
Best Local Similarity 19.4%; Pred. No. 16;
Matches 110; Conservative 67; Mismatches 195; Indels 195; Gaps 22;

QY 758 PFLGALTMPAVRTEALTSGLAVIDLFLQMSAFVALLSLDSKQBSRLDCCVCKP-- 815
Db 79 FMMSLDAVALNVFAGTNTLATV-----VCAFASDLYLGRYATVAG-----CVSTFI 127

QY 816 -----QELPPPGQGGELLLGPFQKAYAPFLHWHITRGVLLFLA----LFGVS 860
Db 128 GMVILTMTAGVPALHPPPCGEGRCIG-----ATRGQAVLGLAFATVAGAG 174

QY 861 LYSMCHISVGLDQELALPKDSYLLDYFLPLNRYFVGAPVYFVTILGNFSSEAGNNAIC 920
Db 175 GIRPCSLPFGADQ-----CPSTVN-SLNCLKNCMSITMGSVRPSVBEQF-----V 362

QY 921 SSAGCNNSFTQKIQVATEPPEQSYLAIPASSWVDDFDLWLTSPSCCRLIYSG----- 973
Db 200 -NSFFNWTYFTLTIIVCASSTAIIVTQSSVSWWVGLAIPALMLASCALFPAGAGLYVRV 258

QY 974 -PNKDKF-----HKYLPWFILNDRPNI---KCPKGLAAVSTSNLTSDDGQVLASRMA 1051
Db 259 RPEGSPFAGVARVAAAFKRSAAPSDADESLFRTHASGVV--SRLPYTDQFRFLDKA 316

QY 1009 -----YHKLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPTTINVEYOY---- 1105
Db 317 AVVVDKASEVGDGHPKPNWRLCSLQQVEEAKCILRVVPMWLTCTI-----V 362

QY 1052 YHKLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPTTINVEYOY---- 1105
Db 363 YVAFPAQNTTYVILQAQSDRHLGGGGAGSFEVP--PGSFTVFPMALAVMILVDRIV 420

QY 1106 -----LTILPEGLFMLSCLVPTPAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAW 1159

Db 421 VPMARLTGREGGITPLQ-----RMGVGNMALSIVLAMLVAAMAEKRER 462
QY 1160 DISYNVS-----LINLVSAGMVSVEFVSHITRSPAISTKPTWLERAKEATIS 1207
Db 463 DLAAQSPNTRGVRSQSAFWLPQLAALGLSFAFNQ-----VSQTEFYTFEPFESMRS 515
QY 1208 M-GSAVFAGVAMTN-LPGILVLGLAKA 1232
Db 516 VAGSVLPFSGLALSSYLSGVLVAVERA 542

RESULT 76

US-11-330-403-7153
; Sequence 7153, Application US/11330403
; GENERAL INFORMATION:

; APPLICANT: Abad, Mark S.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53629)B
; CURRENT APPLICATION NUMBER: US/11/330,403
; CURRENT FILING DATE: 2006-01-12
; NUMBER OF SEQ ID NOS: 19250

; SEQ ID NO 7153

; LENGTH: 618

; TYPE: PRT

; ORGANISM: Oryza sativa (japonica cultivar-group)

US-11-330-403-7153

Query Match 1.5%; Score 103.5; DB 6; Length 618;
Best Local Similarity 19.4%; Pred. No. 16;
Matches 110; Conservative 67; Mismatches 195; Indels 195; Gaps 22;

QY 758 PFLGALTMPAVRTEALTSGLAVIDLFLQMSAFVALLSLDSKQBSRLDCCVCKP-- 815
Db 79 FMMSLDAVALNVFAGTNTLATV-----VCAFASDLYLGRYATVAG-----CVSTFI 127

QY 816 -----QELPPPGQGGELLLGPFQKAYAPFLHWHITRGVLLFLA----LFGVS 860
Db 128 GMVILTMTAGVPALHPPPCGEGRCIG-----ATRGQAVLGLAFATVAGAG 174

QY 861 LYSMCHISVGLDQELALPKDSYLLDYFLPLNRYFVGAPVYFVTILGNFSSEAGNNAIC 920
Db 175 GIRPCSLPFGADQ-----CPSTVN-SLNCLKNCMSITMGSVRPSVBEQF----- 973

QY 921 SSAGCNNSFTQKIQVATEPPEQSYLAIPASSWVDDFDLWLTSPSCCRLIYSG----- 973
Db 200 -NSFFNWTYFTLTIIVCASSTAIIVTQSSVSWWVGLAIPALMLASCALFPAGAGLYVRV 258

QY 974 -PNKDKF-----HKYLPWFILNDRPNI---KCPKGLAAVSTSNLTSDDGQVLASRMA 1051
Db 259 RPEGSPFAGVARVAAAFKRSAAPSDADESLFRTHASGVV--SRLPYTDQFRFLDKA 316

QY 1009 -----YHKLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPTTINVEYOY---- 1105
Db 317 AVVVDKASEVGDGHPKPNWRLCSLQQVEEAKCILRVVPMWLTCTI-----V 362

QY 1052 YHKLKNSQDYT--EALRAARELANITADLRKVPCTDPAPFVPTTINVEYOY---- 1105
Db 363 YVAFPAQNTTYVILQAQSDRHLGGGGAGSFEVP--PGSFTVFPMALAVMILVDRIV 420

QY 1106 -----LTILPEGLFMLSCLVPTPAVSCLLGLDLRSGLNLLSIVMILVDTVGFMAW 1159

Db 421 VPMARLTGREGGITPLQ-----RMGVGNMALSIVLAMLVAAMAEKRER 462
QY 1160 DISYNVS-----LINLVSAGMVSVEFVSHITRSPAISTKPTWLERAKEATIS 1207
Db 463 DLAAQSPNTRGVRSQSAFWLPQLAALGLSFAFNQ-----VSQTEFYTFEPFESMRS 515

QY 1208 M-GSAVFAGVAMTN-LPGILVLGLAKA 1232
Db 516 VAGSVLPFSGLALSSYLSGVLVAVERA 542


```
Query Match 1.5%; Score 103.5; DB 6; Length 1136;
Best Local Similarity 19.1%; Pred. No. 39;
Matches 113; Conservative 84; Mismatches 190; Indels 205; Gaps 27;

QY 749 LCSLSEA--ICFPLGAL-----TPPAVTEFALTSGLAVIDLFLQMSAF 791
DB 602 LLSLSAFPIFFFLAALVCLTMTWVEQRTQGTAKAFYSNG-SILKLYLVGSA 660
QY 792 VALLSLDSKROBASRLDCCVKPOELPPPGQEGGLLGF-----FKAYA 837
DB 661 SVI-----GVLGILIGFPPNIIFNAYKSMYEMPSV 693
QY 838 PFLLEWITRGVVLFLFLAFVGL--YSWCHISVGLDQELALPKDSYLLDYFLFNRYPE 895
DB 694 DIGFTMSYS--LLALFVALFCTTFTAYVACRAELRANAATMRPKAPKIGKIFLERISF 751
QY 896 VCAPYVF--VTTLGVNFSSEAGMNAICSSAGC-----NNPSTQKIOY----- 936
DB 752 IWRGNFTSKVTARNLFRYKQRMMLTVLGVAGCTALILTGFLRNSISDIACQYGOIMK 811
QY 937 --ATBPPEQSYLAIPASSWVDDFIDWLTSPSSCRLYISGPNKDFCPTSVNSLCLNKM 994
DB 812 YDAAIYQDMS--APPAK--ETFDIEMDSNI-----KSLAMSQTN-IETVKSQ 857
QY 995 SITMGSRVPSVEQFHLYLPWFNLDRPNIKCPKGGAAYSTSVNLSDGOVLASRRFMAXH- 1053
DB 858 SAQTTSI-----IVPKNLNLPNYIVLRD--RASHTTEKLTDGAIITEKLAKLFD 906
QY 1054 -KP-----LKNSODYTEALRAARELANITADLRKVPGTDPAPFVPPYITINVFYEOYLT 1107
DB 907 VKPGDTITVKNANDKPOI-----KVSATENYAHYIMTKAYYQOVFK 951
QY 1108 ILPE--GLFML-----SLCLVPTFA--VSCLLGLDLRSLNLLIS 1144
DB 952 EKPSYNLDLMLKOTSEKVESDPAEKLTDKALINVTFSNNVSSL-----NETLDSLN 1005
QY 1145 IVMILVDTGVFMALNDISNVLNLSVAVGMSVEFVSHITRSPAISTKPTWLRARKEA 1204
DB 1006 IIVVLIITSAAALLAFVLYNLN-INV-----SERIRE- 1037
QY 1205 TISMSAVPAGVAMTNLPGILVLGLAKAQLIQIFPFRNLNLTILGLLHGLV 1256
DB 1038 -----LSTIKVLGYPKE-VTMYYVRENIILTMGIAAGFI 1072

RESULT 79
US-60-752-355-4330
; Sequence 4330, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2003-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4330
; LENGTH: 463
; TYPE: PRF
; ORGANISM: Pseudomonas putida KTR2440
US-60-752-355-4330

Query Match 1.5%; Score 103; DB 8; Length 463;
Best Local Similarity 20.0%; Pred. No. 11;
Matches 130; Conservative 66; Mismatches 183; Indels 270; Gaps 32;

QY 203 LNFQGDGTG---GLAPLDITF-----HLLPQGAVGSGIQPLNEGVARCNESQDDVATCS 255
DB 1 MNTVSGDGNLAQKPRHVTMLSIAGIIGAGLFGVSG----- 37
QY 256 CODCAASCPAIARPAQALDSTFVLGMPGSLVLIILCSVFVAVVTTLLGVFRVAPARDKSK 315
DB 38 -----HAIAAGPATIISYFVAGTL---VVLVNRMLGEMAV----- 70
```

```
QY 316 MVDPKKGTSLSDKLSFSTHTLIGQFQCGWGTWVASWPLTILVLSVIPVVALAAGLVFTTEL 375
DB 71 -AHPDTG-----SESTVA--DQAIRGWAGTIGLWYFWFWLVIPIEALAAGHVLNAM 120
QY 376 TTDPVELWSAPNSQARSEKAFHQHGFPPFRFTNQVILTPAPRSSRYDSLLGPKNFSGI 435
DB 121 FPQ-VDSW-----IFALASVLLA----- 138
QY 436 LDLDLLELLELLEQLRLHLQVWSPQAQRNISLDQICYPAPLNDPNTSLYDCCINSLLQYFQ 495
DB 139 -----GTLNF-----SVAKYGE 150
QY 496 NNRITLLLTANQTL-----MGQTSQVWDKDHPLYCANAPLTFKDGATALALSCMADYG--AP 549
DB 151 PEFWPAILKVTAILGFIGLGPAAALMDW-----LPNREVSGLS-GLMAYGGPAP 198
QY 550 -----VPPFL-----AIGGYKGKDYSE--AEAL-----IMTF-----SL 576
DB 199 KQWSAVVGAFTVMFSPFICTEAVTIAASESSDPSRNIATKATRSVIRISTFYILSIFVII 258
QY 577 NNPAGDPRLAQAKLBERAFLEEMR---AFQRRMAGMFQVTFTAERSLEDEINRTAED 632
DB 259 SVVPWNPQLAVVGSYQRA--LEIMNIPNAAFMVDLVVLVAVTSCMNSSI----- 306
QY 633 LPIFATSYIVIFLYISLAL-----GSYSWSRVMVDSKATLGLGGVAVVLGAVMAAMGFF 687
DB 307 -----YIARMFSLAKRGDAPAFLNKTSKGVPPRAAVFG-----STLIGAAIAVLNYP 355
QY 688 SYLGRSLVILQVPPVLVSVGADNIFIP-VLEYQRLPRRPGEPREHVHIGRALGRVAPS 746
DB 356 APKG-----VPEFLASSGATALLVYVMAISQLRMR-----RRLRENT 396
QY 747 M-----LLCSLSEACFPL-GALTPM---PAVRT--FALTSGLAVIDLDEL 785
DB 397 LKRFMWLPYLTWAVIIFIAGALAVMMYTPHRAEVSSTLGLAIVISPL 445

RESULT 80
US-10-703-799B-246
; Sequence 246, Application US/10703799B
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauer, Gregor
; APPLICANT: Lee, Heung-Shick
; APPLICANT: Kim, Hyung-Joon
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING STRESS,
; FILE REFERENCE: BGI-124CPCN
; CURRENT APPLICATION NUMBER: US/10/703,799B
; CURRENT FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: 09/603,208
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142692
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 60/151214
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: DE 19930429.7
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931413.6
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931457.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931541.8
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19932209.0
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: DE 19932230.9
```


	Best Local Similarity	26.8*	Pred. No. 15;	Mismatches	85;	Indels	51;	Gaps	12;
	Matches	63;	Conservative	36;					
QY	1088	PAFEVPPYTINVFYEQLYLILPEG-----LFMLS-----LCLVPTP-AVSCLLALGL	1133	:	:	:	:	:	:
Db	32	PSFHAPQHFSVVVLQACVHLPEGTQNYFESYLAVASVPFMLCLVANELLNVRVAHI	91	:	:	:	:	:	:
QY	1134	DLRSGILLNLISVMIL-----VDTPGMAWDLSYNALSLINLSVAGMSVEFVSHI---	1185	:	:	:	:	:	:
Db	92	RVLASLTVALIAFNVIITALKVDTI-----SSWTRGFPAVTICMVLGSASTVSSSIYGM	147	:	:	:	:	:	:
QY	1186	TRSPAISTKPTWLERAKENIIS---MGSANPVAGVAMTNLPGIL-VLGAKAQLIQIPFFR	1241	:	:	:	:	:	:
Db	148	TGSPFM-----RNSEQALISGGMGTSVAVASLDLAASSDVNRNSALA-----FFLT	194	:	:	:	:	:	:
QY	1242	INLIITLLGLHGLVFLPVILSYVGPDPVNPALAL-----EKKRAEEAAVAVMVAS	1291	:	:	:	:	:	:
Db	195	ATVPLV---LCMWGLYLLSLRELYAYMYRPVLAHVFSGEELPODLSAPSVAS	246	:	:	:	:	:	:

```

RESULT 83
US-60-742-219-208
; Sequence 208, Application US/60742219
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John Verner
; APPLICANT: Bradley, Walter Edward
; APPLICANT: Paquin, Bruno
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Little, Randall David
; APPLICANT: Keith, Tim
; APPLICANT: Cousineau, Johanne
; APPLICANT: Erdewegh, Paul Van
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Gene Map of the Human Genes Associated with Psoriasis
; FILE REFERENCE: 306522-2000
; CURRENT APPLICATION NUMBER: US/60742,219
; CURRENT FILING DATE: 2005-12-05
; NUMBER OF SEQ ID NOS: 7303
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 208
; LENGTH: 752
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-742-219-208

```

Query Match	1.5%	Score 102.5;	DB 8;	Length 752;
Best Local Similarity	20.0%;	Pred. No. 25;		
Matches 100;	Conservative 61;	Mismatches 163;	Indels 177;	Gaps 26;

Qy	30	PGYCAFDECGKPELSGLMTLSNVSLCLNTPARKITGDHLLILLOKICPR-LYTGNTQ	88
Db	15	PGLADSA PCPNVNISGGTFTLSH-----GWAPGS-----LLTYS CQGLYPSASR	62
Qy	89	ACCSAKOLVSLASLSITKALL--TRCPA-----CSDNFV	121
Db	63	LCKSSGQWQTPGATRSLSKAVCKPVRCPAPVSPENGITPRLGSYPVGNNVSFECEGPI	122
Qy	122	NLHCH-NTCSPNQSLF INVTRVAOLGAGOLP----AVVAYEAFQHSFABQSYDSCSRVR	176
Db	123	LRGSPVRQCRPN-GMWDEGATVCDNGAGHCNPGISLGAVRTGFRFGHGD-----KVR	174
Qy	177	VPAAMTAVGTMCGVYGSA LCNAQRWLNFOGDTGNGLAPLDTITPHLLEPGQAVSGTGPL	236
Db	175	YRCSSNLV-----LTGSSRRECO-----GNG-----VMSGTEPI	203
Qy	237	NEGVARCNESOG---DDVA-----TCSCQDCAASCRAIARPAQLDSTFYLGMPGSLVLI	288
Db	204	-----CRQPYSDVPEDDPAALGTSFSHMLGATNPTQKTESLGRKIQI-QRSGHNLV	256
Qy	289	IILCSFVAVVTIILVGRVAPARDKSRWDPKGTSLSDKLSF--SPTHLLGQPFQFGWGT	346

[illegible]

	Query Match	1.5%; Score 102.5; DB 8; Length 779;
	Best Local Similarity	20.6%; Pred. No. 27;
	Matches 139; Conservative	81; Mismatches 244; Indels 211; Gaps 33
<hr/>		
Qy	683	AMGFPSYLGIRSSAVILQVPPFLVLSVCADNIFIVLEBYQLPRPGSPREVIHGALGR 742
Dd	20	ALGUVAGVAIAUALI-VAGYLREV-----LAASFGTPRMVEVGRAVOE 64
Qy	743	VAPSMILCSLSEACFFLGALTMPMAVRTFALTISGLAVILDPLLOM----- 788
Dd	65	GAAAYLR-----RQFTTLAGFVVVIPVLLLLPAENTGAKIGRSI 104
<hr/>		
Qy	789	-----SAFALLSLDSKQRASRLDCCCVKPQELPPGQGE-----GLLIG 830
Dd	105	FPVVGAIIFSALVGVFGMSLATRANTRTAAAAAMTR-----GERAAVRIFAFTGGVVG 155
<hr/>		
Qy	831	PFQ-----KAVAPLLHWITRGVVLILLFLAFGVSLYSMCHISVLGD--- 872
Dd	156	MFTVGLGLGAAAVALVFRDTAPOVLBGFGGAALLAMPFRVGGGIPTKA-ADVGA DLVG 214
<hr/>		
Qy	873	-QELALPKDSYLLDYFLNLRYFEVGPAPVYVTTLGLYNFSSEAGNAICSSA----- 923
Dd	215	KVEQGIPED-----DPRNAATIADNVGNDVGCAGMAADI.PESYAVTLVA A 260
<hr/>		
Qy	924	---GCNNFPSQTQYATEPPEQSYLEIPASSWDDFDID--WLTPSSCRLYISGNKKDK 978
Dd	261	LILGVFAFG----ERGLVFP-----LLIPAVGVVTVAVIGIPAVS PRAGDRTGMSAINRRGF 311
<hr/>		
Qy	979	FCPSTVNSLNCLKNCMSITWGSVPSPVEQFHKLWPFLNDRPNIKCPKGGLAAYSTSVNL 103
Dd	312	FISAWSAI-----GWAVSL-----LYLPSTPADPFGMGQ-----STQSGNPRV 351
<hr/>		
Qy	1039	TSDQVLAASFMAHYHKLKNSQDYTEALARAREELA-ANITADLRKV-PGTDPAPFEPVPT 1091
Dd	352	IAIGA VLIGVLAANA IQLLTCGYFTETGRRPVVDIAKASLTGPA TNILAGI VGLSES AVYS 411
<hr/>		
Qy	1097	---ITNVFYBOYLTILPEGFLMLSLCLVPTPAVSCLLGLDLRSGLLNLLSI VMILTAVTV 115
Dd	412	SLLIGA IIFGAYL--LGSGSVTI AL-----FAVA--LAG-----TG LLTTVGV I--VSMDFP 457

```
QY 1154 GFMAWLDISYNA-----VSLNLVSAVGMSEFVSHITRSPAIST-----1193
Db 458 G-----PVDNNAQIAEMSGMDAGAAIITSLDVGNTTK---AITKGIAIATAVLAAS 509
QY 1194 -----KPTWLERAKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFFRNLNL 1245
Db 510 ALFGSGTDTVTTLARDAGALPEAGRGIVGGLNTA-YPDALV-GLITGASV-VFLFS-GLA 565
QY 1246 ITLLGLLHGLVFLPV 1260
Db 566 INAVGRAAGRVVLEV 580

RESULT 85
PCT-US06-07642-14490
; Sequence 14490, Application PC/TUS0607642
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: MATHUR, Eric J.
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564462014340
; CURRENT APPLICATION NUMBER: PCT/US06/07642
; CURRENT FILING DATE: 2006-03-03
; NUMBER OF SEQ ID NOS: 26898
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14490
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Unknown
; OTHER INFORMATION: Obtained from environmental sample
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (35)...(133)
; OTHER INFORMATION: MMPL family
; NAME/KEY: misc feature
; LOCATION: (1)..(304)
; OTHER INFORMATION: Xaa is any amino acid
PCT-US06-07642-14490

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FQVTFTAERSLEDEINRTTAEDLPF---ATSVIVIVELYISLALGSSYSSWRVMDSKAT 667
Db 65 FQVLLAGTATINDTIKSKMSDMQKFVKVLIIAIFLFLPK-----107
QY 668 LGLGGVAVVL-GAVMAAMGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIISTISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEAI---CPFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELPADVLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244

RESULT 86
US-60-658-984A-14490
; Sequence 14490, Application US/60658984A
; GENERAL INFORMATION:
; APPLICANT: MATHUR, Eric
; TITLE OF INVENTION: NUCLEIC ACIDS AND PROTEINS AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 564463014300/D2170-P1
; CURRENT APPLICATION NUMBER: US/60/658,984A
; CURRENT FILING DATE: 2005-03-04

Query Match 1.5%; Score 102; DB 1; Length 304;
Best Local Similarity 24.0%; Pred. No. 7;
Matches 49; Conservative 39; Mismatches 78; Indels 38; Gaps 10;

QY 611 FQVTFTAERSLEDEINRTTAEDLPF---ATSVIVIVELYISLALGSSYSSWRVMDSKAT 667
Db 65 FQVLLAGTATINDTIKSKMSDMQKFVKVLIIAIFLFLPK-----107
QY 668 LGLGGVAVVL-GAVMAAMGFFSYLGIRSS--LVILQVVP--FLVLSVGAD--NIFIFVLE 720
Db 108 -RLSGVILPLFTVILAIISTISLMAIFSPITVTVQILPSPFLAITIGASVHLLAIFYKD 166
QY 721 YQRLPRPCEPREVHIGRALGRVAPSMLLCSLSEAI---CPFLGALTPMPAVRTFALTSG 777
Db 167 YDF-----SKDKKASLKYALGHSLAIVMTSLTTAAGLWSFSFSELPADVLGKFA-SSG 220
QY 778 LAVILDF-LLOMSAFVALLSLDSK 800
Db 221 ILLSLMFTLILLPALIATINIKPK 244

RESULT 87
US-60-752-355-22441
; Sequence 22441, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 22441
; LENGTH: 569
; TYPE: PRT
; ORGANISM: Wolbachia endosymbiont of Drosophila ananassae
US-60-752-355-22441

Query Match 1.5%; Score 102; DB 8; Length 569;
Best Local Similarity 20.7%; Pred. No. 18;
Matches 111; Conservative 68; Mismatches 146; Indels 210; Gaps 27;

QY 276 FYLGQMPGSLVLIILCSFVAVVTILLVGRVAPARDKSMVDPKGTSLSDKLSFSTHT 335
Db 9 FYYIKPNLSYFVIAFIAVLFSALTILFG-----RGLSNIID--SGT-----EHDFTTKL 56
QY 336 LLGQFFQGGTWTWASWPLILVLSVPVVALAAGLVFTTELTTDPVLMAPSQAQSEKA 395
Db 57 L-----VAIIIVLAISLTA-----FTRLVYFIGI-----GSEKV 84
QY 396 FHDQHFQPPFTNQVILTPAPNSVRYDSILLGPKNFSGILDLLLELLELQERLHLQ 455
Db 85 I-----ARIVYDL-----YSSITDL-----99
QY 456 VMSPEAQRNISLQDICYAFINPDNTSLYDCCINSLLYQFQNNRTLLLTANQTLMGQTSQ 515
```

```
QY 1154 GFMAWLDISYNA-----VSLNLVSAVGMSEFVSHITRSPAIST-----1193
Db 458 G-----PVDNNAQIAEMSGMDAGAAIITSLDVGNTTK---AITKGIAIATAVLAAS 509
QY 1194 -----KPTWLERAKEATISMGSAVFAVAMTNLPGILVLGLAKAQLIQIFFRNLNL 1245
Db 510 ALFGSGTDTVTTLARDAGALPEAGRGIVGGLNTA-YPDALV-GLITGASV-VFLFS-GLA 565
QY 1246 ITLLGLLHGLVFLPV 1260
Db 566 INAVGRAAGRVVLEV 580
```

Db 100 --OPSEFFGVQDVISALIT--DTSVLQSIINSSL-----LITLRFVILIGSVAM 147
Qy 516 VDKOHFLYCANAPITFKDGTALALSCMADYCAPVPPFLAIGYKGYKDYSEBALIMTFS 575
Db 148 L-----LY-TNQLT-----AYAAIIPIL-----LIIVTS 172
Qy 576 LNNYPAGDPRLAQAKLWEBA-FLEENRAPQRRMAGMPQVTTAERSLEDEINRTT-AEDL 633
Db 173 LGKVRSHARFAQDKLSELASFSE-----NFRSIVITIKSPVLEENKTRPKYVL 222
Qy 634 PIPATSYIIVFLYISLALSGYSWSRWVDSKATLGLGGVAVVLGAVMAAMGFFSYLGR 693
Db 223 NSVSKSYVLLV-----RAIL-----VTLVITCVIGSLVLLFFPGIK 260
Qy 694 ---SSLVILQVPPFLV---LSVGA-----DNIFIP-----VLEYORLPRRGPBR 732
Db 261 EVLSNNITIGELSSFFVVSALAGAINNLSIDLQRLGIVERLFEFNKMSIIVDP 320
Qy 733 EVHIGRALGRVAPSMMLCSLSBAI-----CFPLGALTTPMPAVR--TFALTSGLA 780
Db 321 D-----PIKICSVQKGISPGVTFVFEVSQSDKPALNNVSFSIEAGQAV 363

RESULT 88

US-10-498-451-1750
; Sequence 1750, Application US/10498451
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Wang, Jian-ru
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunding
; APPLICANT: Ghosh, Malabika
; APPLICANT: Xue, Aidong J.
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Duntui
; APPLICANT: Goodrich, Ryle W.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Weng, Gezhi
; APPLICANT: Xu, Chongjun
; APPLICANT: Boyle, Bryan J
; APPLICANT: Drmanac, Radoje T
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 820/PCT
; CURRENT FILING DATE: 2004-06-09
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/339,739
; PRIOR FILING DATE: 2002-04-22
; PRIOR APPLICATION NUMBER: US 60/339,453
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/365,384
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/365,091
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: US 60/372,615
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 60/376,045
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 60/372,381
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: US 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 09/552,317
; PRIOR FILING DATE: 2000-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 3134
; SOFTWARE: pt_FL_genes Version 6.0
; SEQ ID NO 1750
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-498-451-1750
Query Match 1.5%; Score 102; DB 6; Length 1964;
Best Local Similarity 19.7%; Pred. No. 1.2e+02;
Matches 174; Conservative 108; Mismatches 257; Indels 346; Gaps 49;
Qy 501 LLLTANQTLMGQTSQVDMKDFLYCANAPLT--FKDGTALALSCMADYCAPVPPFLAIG 558
Db 875 LLPQNTCLMYLLEQIDMHGFGCGTAATSPLTAVFSLSRSLLAAALL-YG-----PCLGA 927
Qy 559 YK GK-----DYSEAEALIMTFSLN-NYPAGDP-----RLAQAKLW-----EPAFLEMRA 602
Db 928 IKTPMPHQHPVFLSVFCGLLVALSYHLSRQSDPTVLSLRSKLFPELEERSLETARA 987
Qy 603 FORRMAGMPQVTTAERSLEDEINRTTARDLPFATSYIIVFLYISLALSGYSWSRWV 662
Db 988 -----EPPDPLDKRQSVREVL-----HSDLMVCVVIATVFAISA-STVFI 1029
Qy 663 DSKATLGLGGVAVVLGAVMAAMGFFSYLGRISLVLQVPPFLVLSV----- 709
Db 1030 ALKSVLGP-----VLVALAGVGFTHYLLPQ---LRKQLPWFCLSQPVLKPLEYSQYEV 1081
Qy 710 -GADNIFIFVLEYQRLPRRGPREVHIGRALGRVAPSMMLCSLS-----EAIQF 758
Db 1082 RGAQVMMPEKLYAGL-----QCVEKYL-----IYPVVLNALTVDATVTVSHDPKYCF 1130
Qy 759 FLGALTTPMPAVRTFALTSGLAVIDPLQMSAPVALLSDSKRQASRLDCCCVKPOEL 818
Db 1131 YCRAL-----LMTVAGLKLRSAP-----CC----- 1151
Qy 819 PPPQGGEGILLGFFQKAYAPFLLHWTIRGWILLFLALFGVSLYSYCHISVGLDQELALP 878
Db 1152 -PPQQ-----YLTLAFTVLLF-----HFDPRLSQG----- 1176
Qy 879 KDSYLLDPLFLNRYFEVGCAPVYFTTGLGNFSSBAGMNAICSSACNNPSTFKIQYAT 938
Db 1177 ---FLLDYPL-----MSLLCSKL-----WDLLYKLRFVL 1202
Qy 939 EPPEOSYLAIAPASSWVDDFIDMLTPSSCCRLYISGPNKDCPSTVNSLNCNKCWSITM 998
Db 1203 -----TYIAPWQITWGSAPFAQFPACATY-----PACV----- 1232
Qy 999 GSVRPSVEQPHKYLPMFLNDRENIKCPKGGLAAYSTSVNLTSQGVLASRFRMAYHKPLKN 1058
Db 1233 -----WAL-----LSGL--FSTPLNPLGSAVF---IMSYARPLKF 1263
Qy 1059 -SQDYTEALR-----AARELAA-----NITADLRKVPCTDPAPFVFPY 1095
Db 1264 WERYKESMSRVHPLFTFLANRLVALSLNSTRVDHNSNTRVLTOLDRNFAG-----DN 1317
Qy 1096 TITNVFYEQLTILPEGLFMLSIC--LV-----PTFAVSCLLIGLDRSLGLMLLSI 1145
Db 1318 NLNSIFYEHLTRSLQH-----TLCGDLVLRGWNVGPQ---DCFVLASLYNALVHLIEV 1369
Qy 1146 VMILVDYTVGFMAWMDISYNAVSLINL-VSAVGMYSVE-----FVSHITR--SFAISTK 1194
Db 1370 GNGLV-----FQLRGLEFRGTVCQQQREVEAITEGVEEDBGCCCEBGLPRVLSFNAFG 1425
Qy 1195 PTWLERAKKATISMGSAVPAVAMTNLPGILVLGLAKQLIQIPFRLMLLTLLGLLHG 1254
Db 1426 QRWL--AWEVITAS--KYVLEGYSISDNN-----ASMLQVFDLR-KILIT----- 1465
Qy 1255 LVPLFVLVSLYCGPDVNPALALEQKAEBAVAAMVASCENHPSRV 1299
Db 1466 -YYVKSIIYY-----VSRPKLEVWLSHSGITAAAL-----RPRV 1499

```
RESULT 89
US-60-781-306-41
; Sequence 41, Application US/60781306
; GENERAL INFORMATION:
; APPLICANT: MESRI, Mehdi et al.
; TITLE OF INVENTION: PROSTATE DISEASE TARGETS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001683
; CURRENT APPLICATION NUMBER: US/60781,306
; CURRENT FILING DATE: 2006-03-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
US-60-781-306-41

Query Match      1.5%; Score 101.5; DB 8; Length 634;
Best Local Similarity 19.3%; Pred. No. 23;
Matches 105; Conservative 74; Mismatches 161; Indels 205; Gaps 28;

QY 522 FLYCANAPLTFKGTALALSCMADYGAPV-----FPFLAIGGYKGYKDYSEAEALIMT 573
Db 72 FLIAALA-----SVNAGLCYAEFGARVFKTGSAYLYTYTVTG-----ELWAPITG 116

QY 574 FSL-NNYPAGDPRLAQAOLWEBAFLEEMRAFORQRMAGMFQVTFABRSLEDEINRTAED 632
Db 117 WNLILSVIGTSSVARA--WSGTFDELL-----SKQIGQFLRTYF-----RMYTGLAE 163

QY 633 LP-----IFATSYIVIFYISLA-----LGSYSSW----- 657
Db 164 YPDFFAVCLILLAGLLSFGVKESAMVNKVTAVNILLVFWVAGVKNVANKISEE 223

QY 658 --SRVVDK-----ATLGLG-----VAVLVGAWAAMGFSY----- 689
Db 224 FLXNISASAREPPSENGTSYAGGPMYGFGTTLAGAAATCPYAFVGFDCIATTGEVRN 283

QY 690 -----LGRSSIVILQVVPFLVLSVGADNIFIFVLEYQRLPRRPGEP-----REVHIGRAL 740
Db 284 PQKAIFIGIVTSLVCFMAYF-----GVSAALTLMPPYLLDEKSPLPVAFYVGVGWGPAK 338

QY 741 GRVAPSMILCSSEALCFPIGALTMPAV-----RT-----FALTS 776
Db 339 YVVAAGS--LCAIISTSL-----LGSIFPMRVIYAMAEDGLLKFCLAQINSKTPPIATLSS 394

QY 777 G-LAVILDFLLQMSAPVALLSLS-----RT-----KQOASRLD 808
Db 395 GAVALMAFLFDLKALVDHMSIGTLMAYSLVAACVLIILRVTSKSSQVMTMLQROGFSMRT 454

QY 809 VCCCKVQBELPPPGQGE--GLLLGFFQ-----KAYAPFLHWTIR-----GVVILLFLA 855
Db 455 LFC---PSSLPTQOSASLSVFLVGLFPLVGLSVLTGYVHAIITREANSLALLALFLV 511

QY 856 LFGVSLYSWCHISVGLDQELA-----LPKDSYLLDYFLF-----NRYFVGV 897
Db 512 LP-VALVLTIRQPNQONQKQVAFWFPFLPFLPAFSLNVIYMLVQLSADTWVRFSIWAIG 570

QY 898 APVYF 902
Db 571 FLIYF 575

RESULT 90
US-10-917-503B-12928
; Sequence 12928, Application US/10917503B
; GENERAL INFORMATION:
; APPLICANT: OTA, TOSHIO
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: NISHIKAWA, TETSUO
; APPLICANT: HAYASHI, KOJI
; APPLICANT: SAITO, KAORU
; APPLICANT: YAMAMOTO, JUNICHI
```

```
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: WAKAMATSU, AI
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: OTSUKI, TETSUJI
; TITLE OF INVENTION: PRIMERS FOR SYNTHESIZING FULL-LENGTH CDNA AND THEIR USE
; FILE REFERENCE: 084335/0123
; CURRENT APPLICATION NUMBER: US/10/917,503B
; CURRENT FILING DATE: 2004-08-13
; PRIOR APPLICATION NUMBER: US/09/629,469
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: JP 1999-248036
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: JP 1999-300253
; PRIOR FILING DATE: 1999-08-27
; PRIOR APPLICATION NUMBER: JP 2000-118776
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-183767
; PRIOR FILING DATE: 2000-05-02
; PRIOR APPLICATION NUMBER: JP 2000-241899
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/159,590
; PRIOR FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: 60/183,322
; PRIOR FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 19025
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 12928
; LENGTH: 397
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-917-503B-12928

Query Match      1.5%; Score 101; DB 6; Length 397;
Best Local Similarity 26.9%; Pred. No. 13;
Matches 57; Conservative 35; Mismatches 74; Indels 46; Gaps 12;

QY 1101 FYEOYLAI--LPEGLFMLSICLVPTF-AVSCLLGLDGLSLNLLSIWML-----VD 1151
Db 24 YFESYLVAASVPS-----MLCLVANFLNVRVAVHVRVILASLTVILAFMVTALVKVD 78

QY 1152 TVGFMALWDISYNAVSLINLVSAVMSVVFVSHI---TRSFATSTKPTWLERAKEATIS- 1207
Db 79 T-----SSWTRGFPAVTIVCNVILSGASTVFSSSYGWTGSFPM-----RNSQALISG 126

QY 1208 --MGSAPVAGVATNLPGL-VLGLAKAQIQLQIFPFLNLLITLGLLHGLVFLPILSY 1264
Db 127 GAMGGTYSVASVSLVLAASDSDVRNSALA-----FFLTATIFLV---LCMGLYLILSLREY 178

QY 1265 VGPDVNPALAL-----BOKRAEBAVAAMVVAS 1291
Db 179 ARYMRPVLAAHVFSGEELPQSLSAPSVAS 210

RESULT 91
US-10-533-519-2249
; Sequence 2249, Application US/10533519
; GENERAL INFORMATION:
; APPLICANT: CLARK,HILARY
; APPLICANT: SCHOENFELD,JILL
; APPLICANT: VANLOOKEREN,MENNO
; APPLICANT: WILLIAMS,P. MICKEY
; APPLICANT: WOOD,WILLIAM I.
; APPLICANT: WU,THOMAS D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT OF IMMUNE
; TITLE OF INVENTION: RELATED DISEASES
; FILE REFERENCE: F1984R1 US
; CURRENT APPLICATION NUMBER: US/10/533,519
; CURRENT FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: PCT/US03/34312
; PRIOR FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: US 60/423,394
; PRIOR FILING DATE: 2002-11-01
```


Qy	1265	VGPDVNPALAL-----EQKRAEEAAVMVAS	1291
Db	257	ARYYMPVLAHHVFGEELPQDLSAPSVAS	288

RESULT 94
US-10-184-615A-360
/ Sequence 260, Application US/10184615A
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin

```

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/
/ FILE REFERENCE: P3430R1C1
/
/ CURRENT APPLICATION NUMBER: US/10/184,615A
/
/ CURRENT FILING DATE: 2002-06-27
/
/ PRIOR APPLICATION NUMBER: 60/059263
/
/ PRIOR FILING DATE: 1997-09-18
/
/ PRIOR APPLICATION NUMBER: 60/059266
/
/ PRIOR FILING DATE: 1997-09-18
/
/ PRIOR APPLICATION NUMBER: 60/062250
/
/ PRIOR FILING DATE: 1997-10-17
/
/ PRIOR APPLICATION NUMBER: 60/063120
/
/ PRIOR FILING DATE: 1997-10-24
/
/ PRIOR APPLICATION NUMBER: 60/063121
/
/ PRIOR FILING DATE: 1997-10-24
/
/ PRIOR APPLICATION NUMBER: 60/063486
/
/ PRIOR FILING DATE: 1997-10-21
/
/ PRIOR APPLICATION NUMBER: 60/063540
/
/ PRIOR FILING DATE: 1997-10-28
/
/ PRIOR APPLICATION NUMBER: 60/063541
/
/ PRIOR FILING DATE: 1997-10-28
/
/ PRIOR APPLICATION NUMBER: 60/063544
/
/ PRIOR FILING DATE: 1997-10-28
/
/ PRIOR APPLICATION NUMBER: 60/063564
/
/ PRIOR FILING DATE: 1997-10-28
/
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/
/ NUMBER OF SEQ ID NOS: 612
/
/ SEQ ID NO 260
/
/ LENGTH: 475
/
/ TYPE: PRT
/
/ ORGANISM: Homo Sapien
/
/ US-10-184-615A-260

```

[illegible]

RESULT 95

```

US-60-732-162-1676
; Sequence 1676, Application US/60732162
; GENERAL INFORMATION:
; APPLICANT: Belouchi, Abdelmajid
; APPLICANT: Raelson, John V
; APPLICANT: Bradley, Walter E
; APPLICANT: Faquin, Bruno
; APPLICANT: Fournier, Helene
; APPLICANT: Nguyen-Huu, Quynh
; APPLICANT: Croteau, Pascal
; APPLICANT: Allard, Rene
; APPLICANT: Debrus, Sophie
; APPLICANT: Eerdewegh, Paul V
; APPLICANT: Little, Randall D
; APPLICANT: Keith, Tim
; APPLICANT: Segal, Jonathan
; TITLE OF INVENTION: Genemap of the Human Genes Associated With Asthma Disease
; FILE REFERENCE: 059908-5010-PR
; CURRENT APPLICATION NUMBER: US/60/732.162
; CURRENT FILING DATE: 2005-11-02
; NUMBER OF SEQ ID NOS: 4417
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1676
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homosapiens
US-60-732-162-1676

```

Query Match	1.5%	Score 101;	DB 8;	Length 475;
Best Local Similarity	26.9%	Prd. No. 17;		
Matches	57;	Conservative	35;	Mismatches 74; Indels 46; Gaps 12;
QY	1101	FVEOYLTI---LPEGLFMLSCLVPTPF-AVSCLLGLGDLDRSGLNLLSLVMIL-----VD	1151	
DB	102	YFESYLAVASTVPS-----MLCLVAVNLLVNVAVHIVRLASLTVLAIFAIVITALVKVD	156	
QY	1152	TVGFMALWDISYNVALSLINLNVSAVGMSEVFSVSHI---TRSFPAISTKPTWLBRAKEARIS-	1207	
DB	157	T-----SSWTRGFPFAVITVCMVILSGASTVFSSSIYGMTGSPFM-----RNSQALISG	204	
QY	1208	--MGSAPFAGVAMTNLPGIL-VLGLAKAQLOIQIFFRNLNLAITLLGLLHGLVFPVLVILSY	1264	
DB	205	GAMGCTVSASVSLVDLAASDVNSGALA-----FFLTATIFLV---LCMGLYLLLSRLEY	256	
QY	1265	VGPDPNPALL-----EQKRAEEVAANVNAS	1291	
DB	257	ARYTMREVLAAHVFSGEELPQDLSAPSPVAS	288	

RESULT 96

```

US-60-752-355-7829
; Sequence 7829, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21 (53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7829
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Oryza sativa
US-60-752-355-7829

```

Query Match 1.5%; Score 101; DB 8; Length 556;
Best Local Similarity 16.9%; Pred. No. 21;
Matches 107; Conservative 92; Mismatches 206; Indels 228; Gaps 27;
QY 499 TLLLLTANQTLMGQT----SQVDKDHFLYCANAPLTFKDGATALASCMADYGAPVPPFL 554


```
Db 71 TWVLASTIVPANGTPEGDKARVQISFLFMSGINTLLQTLVGRTRLPVTVMNASP-AFVVPVL 129
Qy 555 AIGGYGKDYSAEALIMTFSLNYPAGDPRL-----AQAKLWEBAFLEEM-----R 601
Db 130 SI-----AKDFEQ-----NNYASSHQRFTHMTWRAIQGALIVASILNMLGFTSIWG 175
Qy 602 AFQRMA-----GHQVTFPTAERLEBINTTARDLPFATSVIVIFLYIS 648
Db 176 AFARFSPVIMTPVVCVGLGLFALGF-----POVGKCVETGLFMLILA-VVQQVYP 227
Qy 649 LALGSYSSSRVWVDSKATLGLGGVAVLGVMAAMGFESYLGIRS-----SLVILQP 701
Db 228 YFHHYHERITLFEKYSLLCIGIVWAPAILTAGAYNHVSLTKTQQCHCRTDKSYLISS 287
Qy 702 VPFLVL-----SVGADNIFIFVLEYQ-----RLPRRPGEPREV 734
Db 288 APWIKIPYPFQGTPIFTAGHSFGMGA-----VLVSAPFSTGNHFAATLALAGATPPASV 343
Qy 735 HIGRALGRVAPSMML-----CSLSEAICFFLGALTPMPAVRTFALTSLGLAVILDFL 785
Db 344 -LSRSVGLGIGIMFEGIFGAPAGSSVSVENIGLLG-LTKVGSRRVIQISTGFMIFPSIF 401
Qy 786 LQMSAFVALLSLDSKQEARLDVCCVKPQELPPGQGGELLLGFFQKAYAPFLHWT 845
Db 402 GKFGAFFA-----SIPLP----- 414
Qy 846 RGWLLFLALFGVSLYSMCHISVGLDQELALPKDSYLLDYFLFLNRYPEVGAPVVFVT 905
Db 415 --IPAAIFCILGIV-----NAVGSYMQFVNKISMRYIIGLSLFGISVPQYF--- 463
Qy 906 LGYNFSSEAGMAICSSACGNFSTFKIQYATEFPQSYLAIPASSWVDDFDMLTPSS 965
Db 464 --HEVTASASTGPARTNAG-----WENDIIN----- 487
Qy 966 CRLYISGNKDKFCEPSTVNSLNLKNCMSITMGSVRPSVEQHKY-----LPW---FLN 1017
Db 488 --TVFASGPT-----VSLIVASILDNTLEFRGYENDRGLEPWFMPFLH 527
Qy 1018 DRPNIKCPKGLAAYSTSVNLTSDDGOVLASRFM 1050
Db 528 RRGKYSDFRND-EFYSFPIRV---HDVIPSREL 556
```

RESULT 97

```
US-11-174-307B-3734
; Sequence 3734, Application US/11174307B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: BROVER, Vyacheslav
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 2750-1601PUS2
; CURRENT APPLICATION NUMBER: US/11/174,307B
; CURRENT FILING DATE: 2005-06-30
; PRIOR APPLICATION NUMBER: 60/583,671
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,781
; PRIOR FILING DATE: 2004-06-30
; PRIOR APPLICATION NUMBER: 60/583,651
; PRIOR FILING DATE: 2004-06-30
; NUMBER OF SEQ ID NOS: 5544
; SEQ ID NO 3734
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Triticum aestivum
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: Pfam Name: PTR2; Pfam Description: POT family
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
```

RESULT 98

```
PCT-US05-46487-19
; Sequence 19, Application PC/TUS0546487
; GENERAL INFORMATION:
; APPLICANT: Sheets, Joel J.
; APPLICANT: Ni, Weiting W.
; APPLICANT: Larrinua, Ignacio M.
```

```
; OTHER INFORMATION: GI Number: 16226701; NR Description: AT3g47960/Tl7F15_170
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 6730698; NR Description: Similar to peptide
; OTHER INFORMATION: transport proteins [Arabidopsis thaliana]
; OTHER INFORMATION: >gi21436191|gb|AA51383.1| putative peptide transporter protein
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933397; NR Description: nitrate transporter
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933400; NR Description: nitrate transporter
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION:
; OTHER INFORMATION: GI Number: 11933414; NR Description: nitrate transporter
; OTHER INFORMATION: GI Number: 11933414; NR Description: nitrate transporter
US-11-174-307B-3734

Query Match 1.5%; Score 101; DB 7; Length 607;
Best Local Similarity 19.6%; Pred. No. 24;
Matches 116; Conservative 72; Mismatches 219; Indels 184; Gaps 22;

Qy 130 SPNQSLFINVTRVAQLGAGQLPAVVAYEAFYQHSFAEQSYDSCSRVRVPAATLA--VGT 187
Db 86 SVNAATLLNVFS---GTSNLTATVFG-----AVUSDY--LGRVTTTAAATMSSFIGM 132
Qy 188 MCGVYGSAL-----CNAQRWLNFGDGTNGGLAPLDITPHLLBEGQAVGSGIQLNBEV 240
Db 133 LILTLTAATHLTHLPACNASKQHCEGPTGSLAAAILVSFPFL---VIGAG-----GI 182
Qy 241 ARCNESQDDVATCSQDCQCAASCPAIAAPQALDS-----TFVLQMPGSLVL 287
Db 183 RPCNLAFGADQ-----FNPHTADGRGRIASFNNVYFTFTVAMLSATVI 227
Qy 288 IILICSV----FAVVTILLVGRVAPARDKSKMDPKKGTSLDSLKLSFTHTLLGQFQG 343
Db 228 IYLSNVANWALGLAVPAALMGLSCA-----VFFWG 257
Qy 344 WGTWVASWPLTILVLSVIPVVALAAGLVFTELTDPVELWSAPN-SQARSEKAFHQHFG 402
Db 258 TRLVYVRPEGSPFTSFAQLVVAARKRHRRARGDAELFDPHQSKLVSKLAYTDQ--- 314
Qy 403 PFRTNQVILTAPNRSYRYSLLILGPKNFSGILDLILLLELLEQLERLHLQVRSPEAQ 462
Db 315 -FACLDKAAVTPD-----DALCIDGKTPADPWRLCTVQQVEYVKCLARIIPWSSGIV 367
Qy 463 RNISLQDI-CYAPLNPDTNSLYDCCINSLLQVFQNNRTLLLTANQTLMGQ----- 512
Db 368 FFIULTQGTIV-----VLQAAQMDRIRSKSSSQIQQSGFVFMQLAL 411
Qy 513 TSQVDMDHFLYCANAPLTFKD-----GTALALSCMADYGAPVFPFLAIGGYKGD 563
Db 412 TMIIPVYDRFVVVPPALRRFTKREGITLQRIQVGLASV-----ATMVVSAAVEQRKKI 466
Qy 564 YSEAE--LIMTFSLNYPAGDPRLAQAQLWEAEFLERAPORRMAGHFWQVFTTAERSL 621
Db 467 GSSMSCFWLVPQQLLAGLSEAFGAIGQIEFFYRQFPENRS-----VAG----- 510
Qy 622 EDEINRTTAEDLPFATSVIVIFLISLALGSYSSSRVWVDSKATLIGGG 672
Db 511 -----ALYFLGFAMASYASGLMVMVVRATRGGG 540
```

```
; APPLICANT: Bevan, Scott A.
; TITLE OF INVENTION: Second Toxin Complex from Xenorhabdus nematophilus
; FILE REFERENCE: 63566PROV
; CURRENT APPLICATION NUMBER: PCT/US05/46487
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1016
; TYPE: PRF
; ORGANISM: Xenorhabdus nematophilus
PCT-US05-46487-19

Query Match      1.5%; Score 101; DB 1; Length 1016;
Best Local Similarity 18.3%; Pred. No. 53;
Matches 162; Conservative 83; Mismatches 229; Indels 410; Gaps 42;

Qy 41 KNPELGSGLMTLSNV-----SCLS-----NTPARKITGDHLLILLOKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGENACITERLIWSGNTPAEK--GNN--LAGQCVVHYDPTGMN 178
Qy 87 TQACSAKQLVSLASISITKALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTRVAQLG 146
Db 179 QTNSIS---LTSI---PLSITQQLK-----DDSEADMHGMD----- 209
Qy 147 AQOLPAVVAYEAFYQHSFAEQSDYSCSRVRVPAATAATLAVGTMCGVYGSALCNAOR----- 201
Db 210 -----ESGWNALAPESFTSVS-----TTDATGIVL-TSTDAGNKQRIADV 251
Qy 202 -----WLNFGQ-----DTGNGLAPLDITFHLLEP----- 225
Db 252 AGLLQGSMLALKGQEQVIVKSLTVSAASQKLREHNGI---VTTYTEPETQRVIGI 307
Qy 226 -----GOAVGSGI-----QPLNEGVARCNEEQ----- 247
Db 308 KTERPSGHAAGEKILQNLRYEYDVGNVLKSTNDAAEITFRWRNQKIVPENTYTYDSLYQL 367
Qy 248 ---GDDVATCSCDCAASCPAIARPQALDSTFYLG-----QMPGSLVLIILCSVFA 296
Db 368 VSVTGREMANIGRQKNQIPALP-----IDNNTYNTYSTRDYDRGGLTRI----- 414
Qy 297 VVTILLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGMTWVASWPLTIL 356
Db 415 -----RHNS-----PITGNNTYNTMTVSDH----- 434
Qy 357 VLSVIPVVALAAGLVFTELTTPD--VELWSAPNSQARSEKAFDHQFGFPFRTNQVIL-- 412
Db 435 -----SNRAVLEELAQDPTQVDMLFTPGGHQTRLVPQDQLFWTDRBELQQVILVN 484
Qy 413 ---TAPNRSSVRYDS-----LLLGPK-----NPSGILDLDLLE 443
Db 485 RENTTPDQEFYRYDADSQRVKTHTKGTGNSQIORTLYLPELEWRTTYSG-----NTLKE 540
Qy 444 LLEL-----QERLRLHQ----- 455
Db 541 FLOVITVSGSGAQVRVLHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQBEY 600
Qy 456 -----VMSPEAQRNISLDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANQT 508
Db 601 YPGGTAVWAARSQSEADYKTVRYSGKERDATGL-----YYGYRYYSQSWTGRWL 650
Qy 509 LMGQTSQVQDKDHFLYCANAPLTFQDGTALALSCWADYCAPVPFPLAIGGYKGD--YSEA 567
Db 651 SYDPAGEVDGLNLFMRNPNPVSFSDGR-----FPGQGVLAWTGKKAYRKA 698
Qy 568 EALIMTFSLNNYPAGDPRPRLAQAKLWEEAFLEBMRAPFORMMAGMFWQFTTAERSLRDEINR 627
Db 699 -----VNITEHLLEGASFD-----TEL-----KLNK 721
Qy 628 TTAEDLPIFATSYIVIFLYISLA--LGSYSSWSRVMVMSKATIGLGGVAV---VLGAVMAA 683
Db 722 ---GLRTFVLGVGVASLGVKAATIAGASPM-----GIVAAIGGFVSGAV--- 763
```

```
Qy 684 MGPF-----SYLGIRSLVILQVVPFLVLSVGADNIP 715
Db 764 MGPFANNISEKIGEVLSYL-TRKRSVPVQVGAFFVVTSLVTSALP 806

RESULT 99
US-11-314-892-19
; Sequence 19, Application US/11314892
; GENERAL INFORMATION:
; APPLICANT: Sheets, Joel J.
; APPLICANT: Ni, Weiting W.
; APPLICANT: Larriua, Ignacio M.
; APPLICANT: Bevan, Scott A.
; TITLE OF INVENTION: Second Toxin Complex from Xenorhabdus nematophilus
; FILE REFERENCE: 63566PROV
; CURRENT APPLICATION NUMBER: US/11/314,892
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1016
; TYPE: PRF
; ORGANISM: Xenorhabdus nematophilus
US-11-314-892-19

Query Match      1.5%; Score 101; DB 6; Length 1016;
Best Local Similarity 18.3%; Pred. No. 53;
Matches 162; Conservative 83; Mismatches 229; Indels 410; Gaps 42;

Qy 41 KNPELGSGLMTLSNV-----SCLS-----NTPARKITGDHLLILLOKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGENACITERLIWSGNTPAEK--GNN--LAGQCVVHYDPTGMN 178
Qy 87 TQACSAKQLVSLASISITKALLTRCPACSDNFVNLHCHNTCSPNQSLFINVTRVAQLG 146
Db 179 QTNSIS---LTSI---PLSITQQLK-----DDSEADMHGMD----- 209
Qy 147 AQOLPAVVAYEAFYQHSFAEQSDYSCSRVRVPAATAATLAVGTMCGVYGSALCNAOR----- 201
Db 210 -----ESGWNALAPESFTSVS-----TTDATGIVL-TSTDAGNKQRIADV 251
Qy 202 -----WLNFGQ-----DTGNGLAPLDITFHLLEP----- 225
Db 252 AGLLQGSMLALKGQEQVIVKSLTVSAASQKLREHNGI---VTTYTEPETQRVIGI 307
Qy 226 -----GOAVGSGI-----QPLNEGVARCNEEQ----- 247
Db 308 KTERPSGHAAGEKILQNLRYEYDVGNVLKSTNDAAEITFRWRNQKIVPENTYTYDSLYQL 367
Qy 248 ---GDDVATCSCDCAASCPAIARPQALDSTFYLG-----QMPGSLVLIILCSVFA 296
Db 368 VSVTGREMANIGRQKNQIPALP-----IDNNTYNTYSTRDYDRGGLTRI----- 414
Qy 297 VVTILLVGRVAPARDKSMVDPKKGTSLSKLSFSTHTLLGQFFQGGMTWVASWPLTIL 356
Db 415 -----RHNS-----PITGNNTYNTMTVSDH----- 434
Qy 357 VLSVIPVVALAAGLVFTELTTPD--VELWSAPNSQARSEKAFDHQFGFPFRTNQVIL-- 412
Db 435 -----SNRAVLEELAQDPTQVDMLFTPGGHQTRLVPQDQLFWTDRBELQQVILVN 484
Qy 413 ---TAPNRSSVRYDS-----LLLGPK-----NPSGILDLDLLE 443
Db 485 RENTTPDQEFYRYDADSQRVKTHTKGTGNSQIORTLYLPELEWRTTYSG-----NTLKE 540
Qy 444 LLEL-----QERLRLHQ----- 455
Db 541 FLOVITVSGSGAQVRVLHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQBEY 600
Qy 456 -----VMSPEAQRNISLDICYAPLNPNTSLYDCCINSLLQYFQNNRTLLLTANQT 508
Db 601 YPGGTAVWAARSQSEADYKTVRYSGKERDATGL-----YYGYRYYSQSWTGRWL 650
```

QY 509 LMGTQSOVDWKDHFLYCANAPLTFKDGATLALSCMADYCAPVFPFLAIGYKGD-YSEA 567
Db 651 SVDPAGEVDGLNLFMRNRPVFSDDGR-----PPGGVLAWIGKKAYRKA 698
QY 568 EALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAPFORMMAGMFQVTFTAERSLEDEINR 627
Db 699 -----VNIITHEHLEQASFD-----TFL-----KLNK 721
QY 628 TTAEDLPIPATSYIVIFLYISLA-LGSYSSWSRVMVDSKATLGLGVAV---VLGAVMAA 683
Db 722 ----GLRTFVLGVGVASLGVAATAGASPM-----GIVGAAGGVFSGAV--- 763
QY 684 MGFF-----SYLGRSSLVILQVVPFLVLSVGADNIF 715
Db 764 MGFFANNISEKIGEVLSYL-TRKRSVPVQVGAFVVTSLVTSALF 806

RESULT 100

US-11-375-551-16
; Sequence 16, Application US/11375551
; GENERAL INFORMATION:
; APPLICANT: Bintrim, Scott
; APPLICANT: Mitchell, Jon
; APPLICANT: Larrinua, Ignacio
; APPLICANT: Apel-Birkhold, Patricia
; APPLICANT: Schafer, Barry
; APPLICANT: Bevan, Scott
; APPLICANT: Young, Scott
; APPLICANT: Guo, Lining
; TITLE OF INVENTION: Xenorhabdus TC Proteins and Genes for Pest Control
; FILE REFERENCE: DAS-105X
; CURRENT APPLICATION NUMBER: US/11/375,551
; CURRENT FILING DATE: 2006-03-14
; PRIOR APPLICATION NUMBER: US/10/753,901
; PRIOR FILING DATE: 2004-01-07
; PRIOR APPLICATION NUMBER: US 60/441,717
; PRIOR FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 16
; LENGTH: 1016
; TYPE: PR1
; ORGANISM: Xenorhabdus nematophilus
US-11-375-551-16

Query Match 1.5%; Score 101; DB 7; Length 1016;
Best Local Similarity 18.3%; Pred.No.53; Mismatches 229; Indels 410; Gaps 42;
Matches 162; Conservative 83

QY 41 KNPGLSGSLMTLSNV-----SCLS-----NTPARKITGDHLILILQKICPRLY--TGNP 86
Db 124 ESDNLPGRLLTITEQVKGENACITERLIWSGNTPAEK--GNN---LAGQCVVHYDPTGN 178
QY 87 TQACCSAKOLVLEASLSITKALLTRCPACSNFVNLHCHNTCSNPQSLFINVTRVAQIG 146
Db 179 QTNISIS---LTSI---PLSITQQLK-----DDSEADWHGMD----- 209
QY 147 AQLPAVAYEAFYQHSFAEQSYDSCSRVRVPAATLAVGTMCGVYGSALCNAQR----- 201
Db 210 -----ESGKNALAPESFTSVS-----TTDATGTVL-TSTDAAGNKRQIAYDV 251
QY 202 -----WLNFOG-----DTGNGLAPLDITFHLLEP----- 225
Db 252 AGLQGSWALXKQEQVIVKSLITYSAASQKREHNGI-----VTTYTPEPTQRTVIGI 307
QY 226 -----GOAVGSGI-----OPLNEGVARCNESQ----- 247
Db 308 KTERPSGHAAGEKILQNLRYEYDVPVGNVLKSTNDABEITFRWRNOKIWPENTVYDLSYOL 367
QY 248 ----GDDVATCSQDCAASCAPAIARQALDSTFYLG-----QMPGSLVLIILCSVPA 296
Db 368 VSVTGREMANIGRQKNQLPALP-----IDNNNTYNSRTYDYDRGNLTRI----- 414

QY 297 VTILLVGRVAPARDKSKMDPKKGTSLSDKLSSTHTLLGQFPQCGWGTWASWPLTIL 356
Db 415 -----RHNS-----PITGNNYTTNMTVSDH----- 434
QY 357 VLSVIPVVALAAGLVFTELTTDP--VELWSAPNSQARSEKAFHQHFGFPFRTNQVIL-- 412
Db 435 -----SNRAVLBELAQDPTQVMDLFTPGHQTRLVPCQDLFWTRDELQQLVILN 484
QY 413 ---TAPNRSSYRYS-----LLGPK-----NFSGLDLDLLE 443
Db 485 RENITPDQEFYRDADSQRVINKTHIKTGNSSQIQRTLYLPLEWRTTYSG---NTLKE 540
QY 444 LLEL-----QERLRHLQ----- 455
Db 541 FLQVITVGESGQAVRVLFHWETGKPADISNDQLRYSYGNLIGSSGLELSDGQIISQEEY 600
QY 456 -----VMSPEAQRNISIQDICYAPLNPDPNTSLYDCCINSLLQYFQNNRTLLLTANQT 508
Db 601 YPYGGTAVWAARSQSEADYKTVYSGKERDATGL-----YYGYRYQSWTGRWL 650
QY 509 LMGTQSOVDWKDHFLYCANAPLTFKDGATLALSCMADYCAPVFPFLAIGYKGD-YSEA 567
Db 651 SVDPAGEVDGLNLFMRNRPVFSDDGR-----PPGGVLAWIGKKAYRKA 698
QY 568 EALIMTFSLNNYPAGDPRLAQAKLWEEAFLEEMRAPFORMMAGMFQVTFTAERSLEDEINR 627
Db 699 -----VNIITHEHLEQASFD-----TFL-----KLNK 721
QY 628 TTAEDLPIPATSYIVIFLYISLA-LGSYSSWSRVMVDSKATLGLGVAV---VLGAVMAA 683
Db 722 ----GLRTFVLGVGVASLGVAATAGASPM-----GIVGAAGGVFSGAV--- 763
QY 684 MGFF-----SYLGRSSLVILQVVPFLVLSVGADNIF 715
Db 764 MGFFANNISEKIGEVLSYL-TRKRSVPVQVGAFVVTSLVTSALF 806

Search completed: April 11, 2006, 00:57:45
Job time : 106 secs

THIS PAGE IS BLANK

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE IS BLANK